

RULES, RIGHTS, AND HEDGES

JOHN SCHWENKLER

Department of Philosophy, University of Illinois Urbana-Champaign

MARSHALL BIERSON

School of Philosophy, The Catholic University of America

One is sometimes, but only sometimes, justified in pursuing a suboptimal course of action due to a concern that, in attempting the ideal course, one might fail to follow through and so make the situation even worse. This paper explains why such hedging is sometimes justified and sometimes not.

The explanation we offer relies on Elizabeth Anscombe's distinction between *reasons* and *logoi*. Reasons are normative considerations that identify something good or bad that an act will secure or avoid, while *logoi* constrain which actions one may consider at all. We argue that a careful unpacking of this distinction, in light of Anscombe's associated notion of natural unintelligibility, explains why hedging is only sometimes justified. First, the intuitively inappropriate cases of hedging all involve treating *logoi* as though they had the normative bearing of reasons. Second, the action-theoretic differences between reasons and *logoi* ground a principled explanation of why one may hedge when considering reasons but not when considering *logoi*.

Having made this argument, we then show how our approach to hedging is superior to the two approaches recently developed by Berislav Marušić and Stephen White. First, we show that their accounts give counter-intuitive results in some cases. Second, we argue that their accounts violate a plausible principle of interpersonal parity (i.e., they sometimes allow us to consider the faults of others while precluding us from considering our own faults). Finally, we argue that our account better explains how intuitions shift as we vary the details of cases.

1. Introduction

Sometimes, it seems wrong to treat the anticipated consequences of one's own moral weakness as a reason not to do what would otherwise be morally best. For example, consider this case from Stephen White:

Contact: John Schwenkler <jlschwenkler@gmail.com>
Marshall Bierson <marshall.bierson@gmail.com>

Pierre’s friend has just revealed an embarrassing secret in the expectation that Pierre will keep it between them. Suppose that Pierre betrays his friend’s confidence on the grounds that he will almost certainly do so eventually—he is bad at keeping secrets—and so he may as well do so now, given some opportunity for him to gain through the revelation. ... Since he has little hope of remaining a loyal friend for long, there’s not much to be gained by forgoing [the] benefits of immediate betrayal. Thus reasons Pierre. (2021: 833–834)

Like White, we find it *obviously* wrong for Pierre to reveal his friend’s secret on the grounds he has in mind: that since he will almost certainly reveal the secret eventually, he may as well do so at an opportune time. Nor does it seem to matter that it’s Pierre himself who stands to benefit from his revealing the secret now—it would be just as wrong to do this because it seems like a good way to cheer someone up or to lighten the mood, even if Pierre thinks that later on he’ll reveal the secret in a way that won’t have such good effects.¹ *That* is not a legitimate justification for spilling the secret. Pierre ought to keep his mouth shut.

These intuitions might seem to support a general prohibition on treating the consequences of one’s anticipated moral weakness or irresolution as a justification for decisions that would be disallowed if one only were “more virtuous, rational, or strong-willed” (White 2021: 836). The idea that there is such a prohibition squares with recent work by Berislav Marušić treating self-prediction as a species of bad faith,² as well as with various forms of deontic possibilism.³ However, other cases pull our intuitions in the opposite direction. Here is one of these, also due to White:

... suppose you recognize it would be nice for you to call your mother to wish her a happy Mother’s Day—though only if you will be kind to her and won’t say things to hurt her feelings. However, given the way your conversations with your mother usually go, you have good reason to predict you’ll end up saying something cruel or cutting once you’re on the phone with her. (2021: 826)

Unlike the case of Pierre, in the Mother’s Day case it seems like you *could* justifiably decide, on these grounds, to send your mother a card instead of calling her. And your rationale for doing this would be just the kind that seemed to be ruled out for Pierre: it is that even though it would be best to call your mother up and

1. As White emphasizes (2021: 834).

2. See Marušić (2012; 2015); and cf. Marušić & Schwenkler (2022). We discuss Marušić’s position at length in §3.1.

3. We discuss the actualism/possibilism debate in §2.3.

remain cordial throughout the call, things are so likely to turn out badly if you try to do this that it seems wise to take a safer route, even if it is clearly non-optimal. In this case, unlike the first one, such a choice seems entirely acceptable.

Let “hedging” be our name for the kind of decision-making that we find acceptable in the Mother’s Day case, but not in the case of Pierre: to *hedge* a decision is to treat one’s own tendency to be irresolute as a reason for choosing a course of action that, if one were better at following through on one’s decisions, it would be wrong for one to choose. To the extent that our intuitions about these cases are appropriate, they raise a puzzle. The puzzle is to explain why hedging is sometimes—but only sometimes—acceptable, and also to explain what distinguishes the cases where it is acceptable from the cases where it is not.

Our paper offers a solution to this puzzle, by way of a distinction drawn by Elizabeth Anscombe in her article “Rules, Rights, and Promises.”⁴ In brief, our solution holds that the cases above are different in that only the second one appeals to what Anscombe calls *reasons*, or independently intelligible outcomes that one can aim to bring about or avoid. By contrast, Pierre’s reasoning in the first case is an appeal to his tendency to contravene what Anscombe calls *logoi*, which are a different kind of normative consideration whose function is to constrain the space of practical deliberation. We develop this solution in detail in §2. Having done this, in §3 we critically discuss two alternative accounts of the reasonability of hedging that have been offered by Marušić and White. Section 4 concludes.

2. Reasons and Logoi

Anscombe’s aim in “Rules, Rights, and Promises” is to shed light on what she calls “logoi,” a genus of which the three kinds of normative consideration listed in her title are species. She argues that normative considerations of these kinds are *naturally unintelligible*, and that in this they differ from ordinary “reasons” that bear on what to do. Whereas Anscombean reasons identify something good or bad that an act will secure or avoid, logoi rather constrain which actions one should consider at all. In this section, we first explain this distinction in more detail and then use it to provide an account of when hedging is reasonable.

4. (1981d: 97–106). Anscombe also has extended discussions of this same distinction in “The Source of Authority in the State” (1981e: 138–146) and “The Question of Linguistic Idealism” (1981c: 118–121). All three papers were originally published within a two-year period, and it seems clear that this was a major focus for Anscombe in the mid 1970s. However, Anscombe had also already developed some of these ideas previously. For instance, her understanding of natural unintelligibility was already worked out at the end of the 1960s in “On Promising and its Justice” (1981b: 14–19).

2.1. *Understanding the Distinction*

For the sake of clarity, in what follows we will no longer use the word “reason” in the ordinary capacious philosophical sense. Instead, we will adopt Anscombe’s technical sense of the word — to be explicated in more detail shortly. Reasons and logoi, then, are the two different types of *normative consideration* that speak for and against various actions.

Taking our inspiration from Anscombe (1981d: 101–102), let’s use the following as prototypical expressions of a reason and of a logos, respectively:

- (1) You can’t move that—it would make the shelf fall down.
- (2) You can’t castle this turn—it would break the rules.

On the surface, (1) and (2) are similar. This can give the impression that just as, in (1), it is to *prevent the shelf from falling* that you are advised not to move the thing, so, in (2), it is to prevent some outcome that you are told not to move your piece. This impression is mistaken, however, as there is a deep difference between the normative considerations that these statements appeal to. We will highlight two aspects of this difference.⁵

First Difference: Logoi are Naturally Unintelligible

For Anscombe, the fundamental difference between reasons and logoi is that reasons are *naturally intelligible* in a way that logoi are not.⁶ Here is how she puts this point, in connection with the nature of promises:

[A] promise contains (perhaps on the face of it just *is*) a future-tense description which the giver then makes come true—or he breaks the promise. The obligation [to keep a promise] is a kind of necessity to make the description come true. But what sort of necessity is that?

We may say: the necessity is one of making the description come true—or being guilty of something. Of what? Of breaking a promise. And what is that? A description which someone gives and which because he

5. For further discussion of how Anscombe understands this difference, and how it fits into her broader ethical work, see Bierson (2024) and Faulconbridge (2021).

6. Anscombe attributes the original discovery of this natural unintelligibility to David Hume; the original discussion can be found in *A Treatise of Human Nature*, §III.II.V. There are also close connections between Anscombean natural unintelligibility and John Rawls’s discussion of institutional practice rules in “Two Concepts of Rules” (1955), as well as Michael Thompson’s discussion of practices in Part III of *Life and Action* (2008). For a different treatment of the topic see Ford (2011: 99–104).

has given it he must make come true or be guilty. Of what? Not to just go on running round in the circle let's try again and say: of an injustice, of a wrong against the one to whom the sign, the description, was given. But what *wrong* was that? The wrong of breaking a promise. . . . We are back in the circle after all. (1981d: 320, ellipsis in original)

This argument is meant to show that understanding what it is to make a promise must already involve understanding a promise's normative status, and vice versa. To promise something is not merely to say that you will do it, but to say that you will do it *in such a way* that if you don't, then you will have done some kind of wrong. And the wrong in question is just that: it is the wrong of *breaking a promise*, of not doing what you promised you would do.⁷

The same point applies to rules. A person does not understand what it is for something to be against the rules unless they understand the kind of normative status that rules possess. If a chess player agrees that a certain move is ruled out, but fails to see how that bears on whether she may make it, that reveals that she does not understand what rules are. The player might understand a lot of descriptive facts, such as that the *Big Book of Chess Rules* lists a certain move in the "Illegal" column. But this is not enough: for the player also needs to understand that the move's being in this column means that it's a move that *can't be made*, in the very sense of "can't" that is at work in the first part of (2). Further, she cannot understand the force of this "can't" except through its connection with the particular wrongness of rule-breaking—where this wrongness has a different bearing on what a chess player should do than, say, the possibility that if she breaks the rule then she'll be caught.

Things are different with reason-giving justifications of the kind exemplified in (1). This case involves two independent events, namely the movement of the thing and the falling of the shelf, that are normatively related in virtue of a causal connection between them: because the shelf's falling would be bad, it would also be bad to do what might bring it about. This means that the badness of moving the shelf is naturally intelligible: it's the badness of doing what *causes* something further, which can be understood on its own terms as bad, to happen.

Now, that's not to say that there can't be bad consequences to violating a logos—as our statement (2) might be supplemented by adding: "... and no one wants to play with a cheater". The point is rather that such consequences supply a different kind of practical consideration than the one that's provided by the logos itself. Here it may help to consider H. L. A. Hart's contrast between fines and fees:

7. For a dissenting account of promises, see chapter 7 of T. M. Scanlon's *What We Owe to Each Other* (2000). For replies to Scanlon that defend this broadly Anscombean view, see Vogler (2016) and Kolodny & Wallace (2003).

A punishment for a crime, such as a fine, is not the same as a tax on a course of conduct, though both involve directions to officials to inflict the same money loss. What differentiates these ideas is that the first involves, as the second does not, an offence or breach of duty in the form of a violation of a rule set up to guide the conduct of ordinary citizens. (2012: 39)

In our terms, the difference between a fine and a fee is that while a fee provides what is only an Anscombean reason not to do something, a fine brings a reason and a logos together.⁸ If, e.g., the law in jurisdiction *A* prohibits dumping chemicals into the river and penalizes violations of this law with a \$1,000 fine, but in jurisdiction *B* anyone is allowed to dump chemicals into the river so long as they pay a \$1,000 fee, then while in both jurisdictions a person has the same reason not to pollute, namely that doing so will cost you \$1,000, the overall normative considerations differ between the two. In jurisdiction *A*, but not *B*, dumping chemicals into the river is a violation of the law.

But what is required to understand this fact? It is not enough to recognize the naturally intelligible consequences of polluting. Rather, to understand a fine for what it is, you must understand that the fine is imposed *because the fined activity is not permitted*. This is an irreducibly normative notion, and so there is no understanding of a fine that does not appeal to the deontic category of permission. This is why, if one had a fully law-abiding and virtuous citizenry, then there would not need to be *any* penalties for the violation of law. The law itself would be sufficient—though there might, however, still need to be fees for purposes of coordination (for instance, in order to internalize various externalities).

Second Difference: Different Roles in Practical Reasoning

Along with the difference in how they are intelligible, reasons and logoi implicate our practical reasoning in very different ways. In deliberating, we appeal to reasons as considerations that weigh for or against various available options. By contrast, we appeal to logoi as considerations that constrain which options are available in the first place.⁹ To see this, let's return to the example of chess.

8. For related discussion of this same passage, see Hyman (2015: 93–94).

9. This is what Anscombe was getting at when she suggests in “Modern Moral Philosophy” that the question of whether to violate a constraint is to be “quite excluded from consideration” (Anscombe 1981a: 40). Contrary to a common misinterpretation, Anscombe is not saying that she refuses to consider consequentialism as a philosophical doctrine. Rather, she is trying to characterize what she takes to be required to understand a logos as such. However, the absolutism defended in that paper is far stronger than anything we will argue for here. For further discussion of this topic, see Bierson (2024).

Suppose you advise a chess player to castle, and she dismisses your suggestion on the grounds that she's already moved her king. At this point, you'd be making a category mistake if you were to go on about the additional benefits of castling. For instance, if you pointed out that not only will castling save a pawn, but it's also the only way to keep her queen from getting skewered two turns down, she would find this all irrelevant. If you pointed out that without castling she will be checkmated a few turns later, your advice would again fall flat. And so on. The reason for this is obvious: while the potential benefits of saving a pawn or preventing checkmate are good in-game reasons for and against making certain moves, this is only if those moves are allowed in the first place. And it's the rules of the game that determine which moves these are.

If logoi were just a type of reason, then in the case just described there would be one reason in favor of castling and another against it, and the player would have to assess which reason is stronger. But that's the wrong way to understand the force of rules.¹⁰ It is not that the chess-player's reasons not to castle are stronger than her reasons not to let herself be checkmated. Nor is it that her reasons not to castle are *lexically* stronger than her reasons to avoid losing—as they would be if, say, someone had made a credible threat to kill her loved ones if she sacrificed a rook.¹¹ Rather, the rules simply don't provide reasons in the relevant sense at all. In dealing with a logos, she is not dealing with a bad *consequence* that might follow from a certain course of action. Instead, she is facing a constrained space of moves available for consideration.¹²

Another way to see this difference is to consider what is causally upstream of these respective considerations. If the state of the shelf gives you a reason not to move some supporting object, it also gives you a reason to tie your shoes so that you don't trip into the supporting object. That is, your (Anscombean) reasons to avoid X-ing are equally reasons to avoid Y-ing, if Y-ing will cause you to X. But rules don't work this way. It is not, as such, against the rules to do what will cause one to break a rule. For example, imagine a variant of checkers in which (i) a stone must always leap directly over an opponent's diagonally adjacent stone, and (ii) one must never leap over pieces on prime numbered turns. In playing this game, you might anticipate that making a certain legal move on turn 20 will force you to make two further moves in virtue of which, three turns later,

10. This problematic way of understanding constraints seems to be implied by much of the work at the intersection of deontology and decision theory. See, for example, Lazar (2015) and Tarsney (2018). Both works are committed to treating logoi just as a special kind of reason (in Anscombe's sense of that term).

11. Michael Huemer expresses this idea as the view that "some normative considerations have categorically greater weight than others, in the sense that considerations of the latter kind, however multiplied, can never outweigh considerations of the former kind" (2010: 3).

12. For interesting discussions of how this gets worked out in an agent's psychology see Anscombe (2006) and chapter 17 of Raimond Gaita's *Good and Evil: An Absolute Conception* (2004).

you will have no alternative but to leap a stone. That this legal move will lead, inevitably, to breaking the rules shows that this is a poorly designed game. And if you think it would be good to avoid this outcome (say, because you don't want your partner to notice how bad you are at game design), then you will have a reason (in Anscombe's sense) not to make the troubling move right now. But this is a totally different kind of consideration from the ordinary way that rules enter our practical deliberation. For the particular move that you consider making in turn 20 is not, as such, against the rules—unless, of course, there is some further rule against making any move that inevitably results in paradox.

Now, while Anscombe was an absolutist about what philosophers like to call “moral” logoi (such as the commands against murder and adultery), such absolutism is not necessitated by the distinctive nature of logoi as such. This is because, while logoi do constrain what actions can be considered in certain circumstances, they need not rule types of action out of consideration no matter what. This is clear in some paradigmatic logos-statements, such as:

(3) You can't take that—it's hers.

If taking the thing is the only way to feed one's starving child, and the person whose it is could do without it, then there might be reason enough to violate this person's right to their property. While this is not how Anscombe wants to understand the command against theft, nevertheless it is a *possible* way to understand how such a principle can allow exceptions.¹³ It would still be a logos, however, as it excludes from consideration the possibility of taking someone else's property *for reasons of certain kinds*. For example, that something is someone else's will rule out taking it *just because you really want it*, just as a move's being against the rules means that you cannot make it *just because you want to win the game*. In other words, the distinctive normative structure of logoi is compatible with their only ruling out action–reason pairs, like [taking something because you want it] or [making a certain move in order to win]. The possibility of doing such things for *such* reasons is excluded from consideration by the governing logoi.

2.2 Using the Distinction to Solve the Puzzle

Now we are ready to answer the question that we raised in §1, of why hedging is reasonable in some contexts and not in others. Put *very* briefly, our diagnosis of

13. Anscombe herself, in her (1981a: 39), follows Thomas Aquinas (from his discussion in *Summa Theologiae* II-II 66.7.2) and insists that such cases just do not involve theft at all, since the original possessor of the property has no genuine right to it. For a sustained defense of this Thomistic view of property, see Marshall Bierson and Tucker Sigourney (2023).

what goes wrong in the case of Pierre is that he is treating an Anscombean logos as if it had the normative force of an Anscombean reason—that is, he regards the wrong of breaking his friend’s trust as consisting just in a bad *outcome* that he should aim to prevent if he can. We will argue, however, that evidence of what may result if one makes a certain decision is relevant to practical deliberation only insofar as there are *reasons*, in the Anscombean sense, to seek or avoid these results. Further, the prospect that one may violate a logos does not, in general, provide such a reason—and the cases where hedging is unreasonable involve treating this prospect as if it did.

Before we develop this solution in detail, let us consider the abstract structure that is present in cases where hedging can seem attractive.¹⁴ In such a case, a person faces three possible courses of action:

- A: The ideal course, where she makes a certain decision and then follows through on it, thereby acting very well.
- B: The non-ideal course, where she makes a different decision than in A and follows through on it, thereby acting worse than she would have in A.
- C: The disastrous course, in which the person makes the same decision as in A but then fails to follow through on this decision, thereby acting worse than she would have in B.

Further, the person is in a position to make either of the decisions in question. Finally, the person has evidence that if she makes the decision in A then she is likely not to follow through on it, whereas if she makes the decision in B then she is likely to carry it out, and that the consequences of the former course of action will be much worse than those of the latter. Our puzzle asks: Why does the anticipated likelihood of realizing the disastrous course of action, conditional on deciding to aim for the ideal one, *only sometimes* make it reasonable to aim for the non-ideal course of action instead?

To understand our answer—that is, to understand why it makes a difference whether the operative considerations are reasons or logoi—we need to identify exactly how the probability of realizing the disastrous course of action makes it reasonable to hedge in a situation like that of the Mother’s Day case. Once that is clear, we can explain why that mechanism won’t generally apply to cases involving logoi.

To start, let’s consider a couple of further cases that are structurally similar to the Mother’s Day case, and which we hope will generate a similar intuition, that in these cases it might be reasonable to hedge:

14. Here we offer a simplified version of the presentation in White (2021: 827–828).

- (a) A friend invites you to attend her wedding. Ideally you'd accept the invitation and attend, but getting there requires taking a plane flight, and you anticipate that you are going to end up bailing on the trip due to your fear of flying.
- (b) You're on a diet that requires you to keep from snacking but believe that, if you don't eat a low fat but nutritionally empty snack now, then you'll be famished later on and buy a candy bar and a bag of potato chips from the vending machine.¹⁵

Both of these seem like cases where the evident risk of irresolution might make it reasonable to hedge. There are two reasons why this is. First, in each of these cases the fact that you are likely not to follow through on a decision seems to undercut your reason for making it, as it weakens the connection between the decision itself, and the outcome that it aims to realize.¹⁶ In (a), part of what counts in favor of accepting the invitation is the good of attending your friend's wedding: therefore, evidence that you won't be willing to get on the airplane is evidence that you won't attain what makes the ideal course ideal, even if you commit yourself to it. Likewise, in (b), what recommends skipping the snack is the cost to your health of unnecessary calories: so if declining to snack now isn't an effective way of avoiding that cost, since it increases the probability that you will snack in a more unhealthy way later on, this means you have less of a reason for refraining. And the same thing is true in the Mother's Day case: since what counts in favor of calling your mother is that it's a way to communicate your love and lead her to have a good birthday, if you believe that you won't do manage to do these things even if you call, that undercuts your rationale for doing so. In each of these cases, the likelihood of your eventual failure makes it less likely that you will achieve what you are after in aiming for the ideal course of action, thus lowering the expected value of embarking on it. Let's call this first kind of consideration one of *disconnection*.

Second, the likelihood of not following through on a decision will sometimes create additional reasons against making it in the first place. In (a), for instance, accepting the invitation will require your friend to reserve you a spot and you to book a flight. Likewise, in (b), the snack you would buy from the vending machine will not only have more calories, but will also cost you some money.

15. White considers case similar to (b), though the reasoning he depicts is obviously evasive and self-serving: it involves abandoning a diet in order to eat "the delicious, extra-cheese pizza that's being served for lunch" (2021: 834). We agree with White that *this* could not be a "responsible attempt by a person to manage the consequences of his vices" (834). However, we anticipate that our case (b) will evoke different intuitions.

16. As Anscombe puts it: "... to say, in one form or another: 'But Q won't happen, even if you do P', or 'but it will happen whether you do P or not' is, in some way, to contradict the intention" to do P so that Q (Anscombe 2001: 36).

Finally, in the Mother's Day case, the offense that you'll cause your mother if you phone her up and then treat her rudely is a significant anticipated cost of phoning. Call this second kind of consideration one of *increased cost*.

The question we now must ask is why considerations of disconnection and increased costs don't implicate logoi in the same way they implicate reasons. As a start, consider these variants on the cases just discussed, which we hope will elicit the opposite intuitions:

- (c) As in (a), only you have already promised your friend to attend the wedding and now you receive an equally attractive invitation for the same dates, which won't require taking a plane flight.
- (d) As in (b), only you are fasting as part of a religious observance.

The norms against breaking a promise, as in case (c), or violating a religious obligation, as in (d), are properly understood as Anscombean logoi. And this explains why they are not undercut by considerations of disconnection in the same way as the Anscombean reasons that were at work in cases (a) and (b). In (c), what prohibits you from accepting this new invitation is not just that this will cause your friend to be disappointed, but that since you have *promised* to come you are *obliged* not to accept conflicting invitations (at least, not without your friend's blessing). Likewise, in (d), the point of your fast is not just to limit your calorie intake, but rather that it follows from a principle—that is, a *rule*—that you take to govern your life. These logoi may allow exceptions: you could break the promise, or eat the snack, if doing so were the only way to save someone's life. But breaking a promise, or violating a religious fast, *because you are likely to do so eventually*, are the kinds of action–reason pairs that these principles clearly prohibit. If promises and rules had the normative force of Anscombean reasons, then the rationales on display here would be perfectly sensible. And now we can see why they are not: what binds you not to accept the new invitation, or eat the harmless snack, is not the threat of some bad *outcomes* you have *reasons* to keep from bringing about, but rather that these are kinds of *action* that are ruled out by *logoi*.¹⁷

The same point applies to considerations of increased cost. In (c), turning down the new invitation would mean missing out on this other event, along with

17. This is how Anscombeans avoid the sort of epicycles that deontologists face when they try to understand logoi in terms of reasons. For instance, if what counts against stealing is some unique badness that attends to acts of theft, this suggests that you should steal on a given occasion if it's the way to prevent two acts of theft of the same kind. To avoid this absurd consequence, deontologists claim that you instead have reason to prevent *your own* stealings. Of course, this implies that present theft is justified if it prevents theft on the part of your future self. Worries like this lead Christa Johnson to propose that we actually have reason to prevent agent-centered and *time-indexed* thefts (2019). Yet what if stealing is the way to prevent your time-traveling future self from going back and stealing two things at this very moment?

the unpleasantness of trying to force yourself onto a flight. Again, however, this cannot possibly be a good reason to break the promise you have made: While clearly it is permissible to miss your friend's wedding for certain reasons, doing so *because you want to do something else* is just the kind of action–reason combination that the obligation to keep your promise rules out. Likewise, in (d), while buying a snack from the vending machine might cost some money, and be worse for your health, than eating the snack right now, these are not legitimate reasons for violating your religious obligation. While it's true that your anticipated moral weakness makes it potentially costly to keep the fast, consideration of such costs is *precluded* (or *preempted*,¹⁸ *trumped*,¹⁹ or *silenced*²⁰) by the rule that requires fasting, and so the anticipated cost provides no justification at all for hedging.

While these last two cases both involve moral constraints, it is worth reiterating that the reason/logos distinction is an action-theoretic concept that can apply even when choosing between morally permissible options. The following case, originally suggested to us in conversation by Eric Marcus, illustrates this well:

- (e) You used to be a smoker, but last year you resolved to quit. Soon you will be meeting an old smoking buddy, and you believe that you're likely to smoke when you are together. Because your friend smokes terrible cigarettes, you consider buying a pack of good ones to bring with you, so that if you do lapse it will be by smoking something good.

In (e), we have the intuition that there would be a kind of practical confusion involved in hedging, even if smoking would not be morally wrong. However, the reason why you should not hedge is not that there is any moral constraint against smoking, but rather that it would mean violating a *rule* you have given yourself not to smoke.²¹ Were the situation different, say because you had quit smoking not by adopting a resolution to quit but rather by calculating that the pleasure of each individual cigarette is outweighed by its costs to your health, then hedging seems much more reasonable, as it's just a matter of factoring disconnection and increased cost into your weighing of the reasons.²²

Now we can extend this account to the case of Pierre. In thinking that the likelihood that he'll share his friend's secret eventually means that there is less of a reason to keep it to himself now, Pierre is acting as if his obligation to his friend

18. See Raz (1999: ch. 2).

19. See Dworkin (2001: "Introduction")

20. See McDowell (1979: 345).

21. We say more about the practical significance of such self-imposed rules at the very end of §3.3.

22. However, see White's (2021: 837–838) discussion of what he calls "agency-supporting acts" for a different sort of reason you might have not to hedge even in such a case.

could be undercut by considerations of disconnection, similar to the norm that dictates limiting your calories in case (b). But fidelity is not like this: what binds Pierre to keep quiet is not that it would be a bad thing if he were to share the secret, but rather that this is *a thing he is not to do*—at least, not just for any old reason. Likewise, in thinking that he might as well spill the secret at an opportune time, Pierre is acting as if the cost of sharing the secret inopportunistically, or the benefit of sharing it when there is something to be gained, could be relevant to the question of whether to share it now. And these are exactly the kinds of considerations that are silenced by the obligation of fidelity: *such* reasons are not candidate reasons for breaking a friend’s trust, even if there’s a risk that he is going to do so regardless.²³

2.3 *The Actualism/Possibilism Debate*

Many readers will be familiar with cases of this sort from the actualism/possibilism debate in ethics (not to be confused with the actualism/possibilism debate in the metaphysics of modality). For those readers, it may prove helpful if we orient our own project against that backdrop.²⁴ Consider the following much-discussed case:

Professor Procrastinate receives an invitation to review a book. He is the best person to review the book, has the time, and so on. The best thing that can happen is that he says yes, and then writes the review when the book arrives. However, suppose it is further the case that were Procrastinate to say yes, he would not in fact get around to writing the review. Not because of incapacity or outside interference or anything like that, but because he would keep on putting the task off. (This has been known to happen.) Thus, although the best that can happen is for Procrastinate to say yes and then write, what *would* in fact happen were he to say yes is that he would not write the review. Moreover, we may suppose, this

23. Though we’ll criticize the details of his position below, there are parts of White’s diagnosis of this case that are very much aligned with the account we’ve just given. For example:

Pierre should not betray his friend for the sake of amusement. He should not even betray his [*sic*] friend in order to cheer someone else up. His interest in getting a good laugh, or in brightening someone’s day, is overridden by his duty not to betray his friend’s trust. The fact that he expects to betray his friend sooner or later does not change this. (2021: 836)

However, as we discuss in detail in §§3.2–3.3, White’s account of this overridingness is subject to counterexamples, and inferior to the one we have given here.

24. Readers unfamiliar with this debate or readers who, like us, are skeptical of the terms in which the debate has developed are welcome to skip this section. Nothing in the rest of the paper rests on what we say here.

latter is the worst that can happen. It would lead to the book not being reviewed at all, or at least to a review being seriously delayed. (Jackson & Pargetter 1986: 235)

Whether the professor should agree to review the book seems to depend on whether the professor's obligations are fixed by what a person *will do* if they make a certain choice (the actualist position) or instead what the person *could do* if they make a certain choice (the possibilist position). Cases like this one have spawned an enormous literature, the details of which we will not enter into here.²⁵ The reason for this is that the terms of this debate are really quite foreign to the sort of Anscombean philosophical psychology that we favor. Along with its questionable determinism about "what *would* in fact happen" were agents to make certain choices, the literature generally presents deontic obligations as a special sort of all-things-considered evaluation of one's first order reasons, rather than as a genus of multifaceted first-order logoi. Where Anscombe thinks our deontic obligations are exhausted by ordinary duties such as the duty to keep our promises, our duties to obey God, our duties to our children, etc., this literature instead is focused on some higher-level duty to act on the full balance of reasons.²⁶ Because of this, the standard terms of the debate don't allow for easy translation into our preferred framework. Nevertheless, reflecting on examples like the one above can help show the philosophical interest of our account.

The crucial thing to bring out is that it matters to the case whether there are any Anscombean logoi that bear on Procrastinate's decision. This becomes clear if we imagine a variant of the case in which Procrastinate has already accepted the invitation to write the review, and then an opportunity comes along to work on something more interesting instead, and which he is more likely to follow through on. We expect the reader will agree that Procrastinate clearly cannot hedge his decision in this case, even if the expected cost-benefit calculation points in favor of abandoning his commitment and taking up the new project. By contrast, if Procrastinate is in a situation where he has no standing obligation at all to write the review, then things look different: in thinking about whether to accept the invitation, it seems wise for Procrastinate to base his decision partly on the likelihood that he'll fail to get it done (a consideration of disconnection) and the inconvenience that this would likely cause (a consideration of increased cost).

Of course, in many real-life cases the deliberative situation of a Procrastinate won't be quite so clearly defined. For example, in Holly Smith's original version of this case,²⁷ the invitation you receive is to provide comments on a paper for a

25. For an overview of this literature, see Timmerman & Cohen (2020).

26. For a detailed analysis of Anscombe's account of practical reasoning, see Schwenkler (2019: ch. 5).

27. See Goldman (1978: 185–186).

graduate student in your department, and the next-best person to provide comments on the paper would do a significantly worse job than you would, with negative consequences for the student's career. In Smith's case, your position as a faculty member in the department might obligate you to supply comments—that is, to say yes to the student and then immediately implement whatever accountability mechanisms are required to ensure you will actually follow through.²⁸ By contrast, in a situation where you are totally free of such logos-type constraints, as if the request comes from a student at a different institution, it seems entirely reasonable for you to hedge your decision in light of your tendency to flake out. The structure of your deliberative situation makes an important difference to how you should reason, just as we saw in our pairs of cases above.

In short, our argument suggests something like a *moderate actualism* as the correct position in the actualism/possibilism debate, though only if the terms of that debate are shifted somewhat. The position is actualist, since it says that deliberating agents can often base their decisions on predictions of what they are actually—or, as we would prefer, likely—going to do. But the actualism is moderate, since it says that there are some situations (and indeed precisely those situations which involve the common deontic obligations) in which considerations of this kind are silenced by Anscombean logoi. Of course, given our doubts about the presuppositions of this debate we don't think our view can be recruited in defense of a position that takes them for granted. Working out these doubts more fully would be a task for another paper.

2.4 *Anticipating an Objection*

The reader might worry that our argument proves too much. For there seem to be cases where the consequences of your expected future failure *do* justify violating a logos. And it might seem that our claim of preclusion or “silencing” is incompatible with this possibility. Let us explain why this isn't so.

Suppose you have promised to attend a close friend's anniversary party, and only later do you learn that the party will be at a Vegas casino, where you might relapse into your chronic gambling addiction, thereby risking your family's economic security. Here, it seems like the risk that you will gamble could be an acceptable reason for breaking your promise. While in the ideal case you would go to the party and not gamble, the significant likelihood that you will end up gambling seems to justify backing out. This suggests that hedging is sometimes appropriate even when it comes to options that are precluded by logoi.

28. In the situation where you are so deeply inclined to procrastination that nothing you do will enable you to follow through, it seems to us that you must resign your position. If you lack the requisite character to fulfill your core contractual obligations, then you have no business keeping that job.

To address this worry, we need to recall a point that we made several times above: that while logoi generally structure deliberation, they need not preclude the consideration of countervailing reasons *no matter what*. That is, while the normative force of a promise precludes the consideration of certain ordinary reasons not to act as one has promised to, such “silencing” is not absolute. And this last case is one where the risk to your family’s finances—in contrast, say, with the opportunity to do something else that would be more fun—could be the right kind of reason to go back on your promise to your friend.

However, we need to be clear about the way in which this happens, as this point will be important to our argument in §3. On our account, the explanation of why, sometimes, you may not hedge—that is, why you may not treat the expected consequences of choosing a certain course of action as a reason to choose otherwise—has nothing to do with whether those consequences would result from your own indecision or moral weakness, rather than in some other way. Rather, the explanation is that the operative logoi preclude considering consequences of that kind *regardless* of their source. If, for example, you should not steal a small thing now in order to keep yourself from stealing a larger thing in the future, then similarly you should not steal a small thing to prevent someone else from stealing a larger one.²⁹ A person’s right to their property precludes the consideration of *such* benefits as a potential justification for taking it, and this is why you are precluded from considering these consequences when they would result from your possible weakness of will. That is, it is not that you are screened off from considering the likelihood of your own moral weakness, but rather that, in this case, the anticipated consequences of that weakness are of a kind that provides no justification at all for the hedged course of action.

To see this more clearly, consider one more case. Most people intuit that it is permissible to take someone’s property if doing so is necessary to save a life—as, e.g., a starving man is permitted to take bread from a bakery without paying. Given this, our account predicts that it will be similarly permissible to steal if doing so is necessary to prevent yourself from doing what will endanger a life. And this is exactly what we find. Imagine someone about to head into the desert to rescue a stranded child. She has no money but enough food and water to get them both back safely, but only if she reserves all the food and water for the

29. Of course, this needn’t entail that you shouldn’t *take* (without permission) a piece of someone’s property in order to prevent another person from doing something worse (say destroying the property)—it’s just that you’d have an obligation to give it back, or at least explain to the person what had happened. Notice, however, that even here the intuition relies on the thought that such taking would not really be theft—perhaps because you can presume consent. For instance, if the owner of the item told you that they don’t want your help and that they would rather have the item destroyed than taken without permission, then the intuition you can take it vanishes.

child, without eating or drinking anything herself—an extremely difficult thing to do, so that it’s likely that she will break down and have something to eat or drink, thereby endangering the child’s life. In this situation, our intuition is that it would be justifiable for the person to take some food from a bakery without paying—a kind of justifiable *logos* violation that trades directly on the consequences of her own anticipated irresolution. We expect that those readers who think you can take another’s property to save the life of a starving child, will also find it appropriate to hedge in such a case.

Put differently, when we follow Anscombe in saying that the anticipated consequences of doing otherwise can sometimes justify breaking a promise or stealing a loaf of bread, we are *not* saying that these consequences make a difference to whether or not you are permitted to predict what you will do if you make certain choices or end up in certain circumstances, or to base your decisions on the consequences of such predictions. Indeed, at a single moment the risk of your anticipated moral weakness might make no difference to one practical question (e.g., whether to spill a friend’s secret), but could be relevant to a different one (e.g., whether to get out of the situation in which you are tempted to do so). It is not that you are screened off from considering the likelihood that you will act poorly if you make one decision rather than another, but rather that certain consequences are simply irrelevant to deliberation of this kind, depending on how it is structured by Anscombean *logoi*.

3. Our Solution and Two Others

We argued in §3 that the explanation of why hedging is only sometimes reasonable is that sometimes the anticipated consequences of one’s irresolution are rendered irrelevant by the *logos*-type norms that structure one’s deliberative standpoint. This solution differs from two other recent attempts to explain when and why it is reasonable to hedge. In this section, we show why our solution should be preferred to these other ones.

3.1 Marušić

A central claim of Berislav Marušić’s *Evidence and Agency* is that hedging is unacceptable to the extent that it involves an *evasion* of one’s responsibility as an agent.³⁰ For Marušić, to evade a decision is to think of what one will do as something that is settled by evidence, rather than by the reasons for and against doing

30. For ancestors of this argument see Marušić (2012; 2013).

various things. More precisely, here is Marušić's official definition of what he calls "epistemic evasion":

... we epistemically evade our decision if and only if we treat a practical question as a theoretical question; that is, if and only if when it is up to us to φ , we settle the question of whether we will φ by considering our evidence about whether we will φ if we decide to φ rather than by considering our practical reasons for and against φ -ing. (2015:112)

According to Marušić, cases like our (c) and (d) both involve evasion of this kind, and this is why they don't permit hedging. That is, in (c) you settle the question of whether you will go to the wedding by considering the evidence of your tendencies, rather than the practical reasons that oblige you to attend. Likewise, in (d) you treat the question of whether you will break the fast as settled by the evidence of how likely you are to do this, and not only by the practical reasons against doing so. And similarly, of course, for the case of Pierre. For Marušić, thinking in this way means evading your responsibility as an agent.

How does this position account for the reasonability of hedging in cases that pull in the opposite direction? Here, Marušić would appeal to a pair of distinctions.³¹ The first distinction is between basing a decision on the *likelihood* that one will follow through on that decision if one makes it, and basing a decision on the *difficulty* of following through on it—where this "difficulty" is a matter of associated transaction costs. If, e.g., I am considering whether to quit my morning coffee habit, then it is appropriate for me to base this decision partly on how quitting "would take a lot of willpower, ... would be unpleasant, ... would interfere with my work, and ... would detract from other commitments that I now have" (Marušić 2015: 130), as these are all practical reasons that speak against doing what it takes to quit. Because of this, if I treat them as reasons to keep drinking coffee then I have not thereby evaded my responsibility, since my decision is not based on any "predictions about the chances that I would succeed in drinking less coffee if I decided to" (130). This is the first way that Marušić allows us to hedge our decisions in light of our anticipated frailty: when the evidence shows that our frailty will make it *difficult*, because costly or inconvenient, for us to follow through on a certain decision, this can be a legitimate reason to decide something else.

Marušić's second distinction is between reasoning concerning *whether or not* to do a certain thing, and reasoning concerning *how to go about* doing something

31. Marušić and Schwenkler give roughly the same account in (2022).

that one has decided to do.³² Again, he holds that it is only in reference to the first kind of reasoning that hedging is an unacceptable form of evasion. To illustrate, consider Marušić's discussion of a habitual gambler who "strategically chooses to stay at home" rather than going to casinos, "because otherwise he would be exposed to temptation ..., and there [would be] a significant chance that he would give in" (Marušić 2015: 133). As Marušić understands this case, the gambler is reasoning *from* the decision not to gamble *to* the further decision to avoid going into casinos, on the grounds that if he goes to a casino he is likely to gamble. Unlike the habitual coffee drinker, the gambler's decision is based partly on predictions about the *likelihood* that he will gamble, conditional on deciding one thing or another, rather than the *difficulty* of following through on certain decisions he might make. However, the gambler's choice to stay at home does not fall within the scope of Marušić's definition of epistemic evasion. That is because the question that the gambler is trying to settle is not *whether he will gamble* but rather *whether he will visit casinos*—and the evidence that, if he does go into a casino, he is likely to gamble when he's there, is clearly a legitimate (because appropriately practical) reason in favor of staying out.³³

In sum, Marušić's response to our puzzle centers on two claims, both concerning situations where it is up to you whether or not to φ :

- Evidence of the *difficulty* of φ -ing, where this is understood in terms of the anticipated costs of following through on a decision to φ , can be a legitimate reason to decide not to φ .
- Evidence of the *likelihood* that, even if you decide to φ , nevertheless you will fail to do so, can be a legitimate reason to keep out of those situations where you appear to be likely to φ , though not a legitimate reason to decide not to φ .

This position yields the right verdicts about cases like Marušić's coffee drinker and habitual gambler. Unfortunately, it is inadequate as a general solution to our puzzle.

The first problem for Marušić's position is that it seems to yield the wrong verdict about some cases. One of these is the Mother's Day case. For there is a plausible version of this case where the reason not to call your mother is not only that it will be *difficult*, in Marušić's sense, to talk to her over the phone without saying anything insulting, but rather that you are too *likely*, even if you decide to call her and refrain from any insults, to end up insulting your mother anyway. If

32. For Marušić, this is the distinction between treating theoretical considerations "as the basis for prediction of what we would do if we decided to do something," and treating such considerations "as the basis for a decision" (2015: 128).

33. For further discussion of Marušić's account of such strategic choice, see Schwenkler (2022).

that is how the case is understood, then Marušić must regard your reasoning as unacceptable. And this seems like the wrong verdict to give.

Marušić might respond to this objection by questioning whether in White's case it really is "up to you" whether to refrain from insulting your mother if you call her, at least in the sense that he gives to this phrase. For Marušić, φ -ing is "entirely up to us" when there is "no gap between φ -ing and continuing to try to φ ", so that "if one fails to φ , one did not try hard enough or long enough" (2015: 19). Therefore, if what explains why you are likely to insult your mother is not that you will not have tried hard enough or long enough to refrain from doing this, but rather that you might do such a thing *even when you are trying your hardest* not to, then the decision not to call your mother will not count as epistemically evasive according to Marušić's account.

Unfortunately for Marušić, this interpretation of the case is not obligatory. Most obviously, suppose your worry is simply that if you are on the phone with your mother, then you will get to the point where you cease *even to try* to be polite, and then will say something cruel for this reason. (Indeed, this is how White himself describes the case: it is one where "you suspect that any decision or resolution to speak kindly to your mother is one you'll ignore or abandon once you are on the phone with her" [2021: 830].) Nevertheless, this seems like a perfectly acceptable reason to send a card instead of calling.

The same point can be made in connection with cases (a) and (b). In each of these cases, what inclines you to hedge is not just the difficulty, in Marušić's sense, of carrying out the ideal decision, nor does the non-ideal course just involve keeping yourself out of situations where you are likely to act unwisely. Rather, these are both cases in which it seems reasonable for you to decide not to do something on the grounds that, if you do decide to do it, you are likely not to follow through—and not only because following through is not wholly "up to you." So Marušić's account yields a counterintuitive verdict about these cases too: hedging in these cases is necessarily unreasonable by his lights, even though it is entirely up to you whether or not you follow through on the ideal course of action. And this does not seem right.

Some readers may question our intuitions about these cases, or our insistence that such doings are "entirely up to you" in Marušić's sense. Rather than belaboring these questions, in §3.3 we will raise two more objections to Marušić's account whose force is independent of these considerations.

3.2 White

We turn now to White's account. In "Self-Prediction in Practical Reasoning," White attempts to solve our puzzle by appeal to a constraint against what he

calls *opportunism* in the choice of non-ideal options, where a choice is opportunistic if it is a way “to take advantage of anticipated wrongdoing or prudence to secure benefits [a person] would have to forego—or avoid costs she would have to bear—if she were more virtuous, rational, or strong-willed” (2021: 836). White motivates this account with the case of Pierre, arguing that the problem with Pierre’s reasoning is that it is a way of “taking advantage” of the fact that he is likely to spill the secret eventually, by treating it as an excuse to do what will be somehow beneficial.

More precisely, White’s proposal is that a consideration can favor choosing a non-ideal option, *B*, over aiming for the ideal course, *A*, only if this consideration falls in the class of what he calls *ideally-undefeated* reasons favoring *B* over the yet worse course of action, *C*—where a reason favoring *B* over *C* is ideally undefeated if and only if it does not also favor *B* over *A* in some way (for this definition see White 2021: 835–836). In the Mother’s Day case, your reason for sending your mother a card is ideally undefeated since it wouldn’t also favor sending a card over calling your mother if you thought you could refrain from insulting her. (Contrast a case where you just didn’t want to spend time on the phone.) The same holds for (a) and (b), where what motivates you to prefer the non-ideal option to the worse one would not also lead you to prefer it over the ideal course of action, assuming that you were sufficiently likely to follow through on the decision to pursue it. (That is: were you not afraid of flying, you would have no reason not to accept the invitation. And if you believed you would be able to stick to your diet the rest of the day, you would not want to eat the nutritionally empty snack.) By contrast, Pierre’s reason for revealing the secret now (*B*) rather than risking the chance that he will reveal it later on (*C*), namely that he will gain something by doing so, is *also* a reason that favors revealing the secret at this time (*B*) over simply remaining steadfast (*A*). As such, this justification for choosing the non-ideal course is ruled out.

As we have explained already, we agree with White that Pierre should not betray his friend’s confidence on the grounds just described. The mere fact that things are likely to turn out better if Pierre decides to spill his friend’s secret at an opportune time than if he decides to remain steadfast, is not an acceptable reason for Pierre to hedge. We doubt, however, whether White’s account provides the right explanation of why this is. One reason for this doubt is that his account yields unintuitive judgments about some other cases, such as the following:

- (f) A billionaire offers to make a massive gift to the charity of your choice if you train for, and complete, a marathon in six months. Unfortunately, you have reason to think that even if you sign up for the race you will fail to do the training necessary to complete it. Further, to register for the marathon you will need to wait in line while your child is giving a piano recital.

Suppose, *arguendo*, that the ideal course of action is to accept the offer, register and train for the marathon, and then complete it, even though this means missing your child's recital. On White's account, the good that would come from attending the recital is not an ideally-undefeated reason to decline the offer, as it favors turning down the invitation over *both* accepting it and then failing to complete the race *and* accepting it and then following through, thereby earning the promised money for charity. (That is, this reason favors *B* over *A* as well as *C*.) As such, White must say that the desire to attend your child's recital will "lend no justificatory support" (2021: 836) to the option of declining the offer. That is to say, if you do decline the offer then the desire to attend the recital should have nothing to do with it.³⁴ And this does not seem right. Even if you are not, in the end, permitted to decline, the idea that you could not even *count* attending the recital as a consideration in favor of such a choice is incredible.

The reader might object that our judgments about this case could be colored by the intuition that it is permissible *simply* to choose attending the recital over accepting the billionaire's offer, entirely independently of whether you are likely to follow through on this agreement. Given that the authors of this paper both have this intuition, we think that the objection has some force. However, this only leaves us with a deeper puzzle, which is that if the choice to decline the billionaire's offer could be justified simply by the desire to attend your child's recital, then it would seem to be even *better* justified by the combination of this desire and the belief that, if you accept the offer and wait in line to register, then

34. Contrast what White says about the case of Pierre:

He should not be entitled to use the fact that he's not going to remain loyal as a justification for pursuing what he would otherwise have to give up. To do that would be to treat his general lack of fidelity as a kind of boon from the standpoint of his desire to amuse others. And that does not seem like an acceptable attitude for him to take toward this flaw in his character. It's therefore illegitimate for him to use the fact that he's bound to betray his friend's trust eventually as a justification for revealing the secret when he calculates that doing so would be to good effect. (2021: 836)

This sounds plausible on its face, but that impression evaporates when it's extended to a case like (*f*). Are you "not entitled" to use the fact that you're unlikely to complete the marathon "as a justification for pursuing what [you] would otherwise have to give up" if you signed up for it, namely the opportunity to attend your child's recital? Would doing so mean treating your flawed character "as a kind of boon from the standpoint of [your] desire" to see your child perform? Of course not! And our account explains the difference. In (*f*), what counts in favor of signing up for the marathon is an Anscombean reason: that if you do so, it will benefit a charity. As such, it is appropriate to base your decision on considerations of disconnection (will you really follow through?) and increased cost (will the time you spend waiting in line, and missing the recital, be wasted?). By contrast, Pierre's deliberation is governed by a *logos*-type obligation not to betray his friend's trust (as White implicitly understands: see the passage quoted in fn 23). This obligation is not absolute: Pierre could spill the secret if this were the way to prevent nuclear war. But any personal advantage he would gain by spilling it is not the kind of consideration that would make it acceptable to do so.

you likely won't manage to complete the race. That is to say, that it will allow you to attend the recital *and* not waste your time training seems to be an even better reason for declining than either of these considerations taken individually. But White's position has the consequence that we are disallowed from summing them together, or indeed from treating the desire to attend the recital as any reason at all to decline. And it is hard to see how this could be.

3.3 *Two Further Objections to Both Accounts*

Intuitions can be untrustworthy, and it's not always clear what features of a given case they are influenced by. There are, however, two other problems for both Marušić's and White's positions that are harder to see a way around.

Objection 1: Both Positions Violate a Principle of Interpersonal Parity

Suppose a friend asks your advice about whom to hire to perform some job. In advising your friend, it seems sensible to consider how reliable the candidates are at doing what they promise—and so, if you know that one of the candidates is likely to flake out, then that will be a good (though perhaps insufficient) reason for advising your friend not to hire them. If, however, you can and should take this stance to other people's tendency to be irresolute, then it seems that you should give similar consideration to your own such tendencies. To do otherwise would involve holding others to a higher standard than you hold yourself. You'd be factoring in the weakness of others, but blissfully ignoring your own.

This is a problem for both Marušić and White, because we can construct cases where their accounts prohibit you from making choices that reflect a consideration of your own likelihood of failing to follow through on a decision, but do permit consideration of the likelihood that others will do the same. Begin with the following case, which will be problematic for Marušić's account but not for White's:

- (g) You have been offered a job as a firefighter, which in your world is a highly-paid position. Unfortunately, you are much less courageous than the person who is likely to take the job if you accept it, and so less likely than they are to do the job well in dangerous situations.

In deciding whether to take the job, it seems important for you to account for your cowardly tendencies in just the same way as you would consider the cowardice of others in deciding whether to offer such a job to them, or advise them

on whether to take it. But Marušić would have us say otherwise. For here is how he characterizes the nature of an agent's perspective on what she herself will do:

... even if their evidence is comparable, there is an asymmetry between the agent's and the observer's view of what the agent will do. What explains this asymmetry ... is that it is the agent herself who is forming a belief about her own future course of action. Because an agent enjoys a practical view of matters that are up to her, she can settle the question of what she will do by considering her practical reasons. And when she settles the question in this way, her view of what she will do is subject to the standards of practical, not epistemic, rationality. (2015: 123–124)

Marušić's description of this agential perspective seems plausible when a choice is being made solipsistically, as it were: *Will I keep the promise? Will I finish the marathon? Will I carry out the duties of this job?* When the question is heard in this way, the sense that there may be a *demand* to do the thing in question—what we would gloss as an Anscombean logos governing the choice—invites the intuition that the answer should be “Yes.” However, this plausibility evaporates in situations where a person has to compare herself to others. In (g), the reason you have against deciding to be a firefighter is not merely that it would be *difficult*, in Marušić's sense, to carry out your duties courageously, nor that it is less than entirely up to you whether or not you will do so. (Again, your worry may be that in the face of danger you will simply *stop trying* to do what you need to, and not that you will fail to do so even when you are trying your hardest.) Rather, the reason has to do with the *consequences* that may follow if you make this decision.³⁵ Given this it seems to be, not only permissible, but outright required that your decision be informed by the likelihood that you'll be irresolute.

Though White does not stress in the same way as Marušić does the asymmetry between first- and third-person perspectives, his rejection of “opportunistic” forms of self-prediction leave him similarly vulnerable to this objection. The following case brings this out:

(h) As in (g), though in addition to being very cowardly you are also slightly less physically able to perform the duties of a firefighter than the person who is likely to accept the job if you decline it.

35. This comes out if we contrast (g) with a variant on it where you have been offered a job as a stunt double instead. Here, you have less reason to decline the job offer than you had to decline the offer to be a firefighter. And that's all because of the importance of what hangs in the balance: if a stunt double is cowardly, then filming will be delayed; but if a firefighter is cowardly, people will die.

In (*h*), the small danger to public safety that arises from your physical limitations might not be enough of a reason to decline the job, if the benefits associated with accepting it are significant enough. If that is so, then the ideal course of action might be the one where you accept the job and carry it out to the best of your physical abilities, including behaving with the necessary courage in the face of danger.

How, though, are you to think about the likelihood that you *won't* do this job as courageously as whoever is next in line? If this likelihood is high enough, then clearly this should count as a consideration in favor of declining the offer. And if that's what you decide, then you will be choosing to pursue the non-ideal course of action on grounds of public safety—that is, in light of the fact that people's lives and well-being would be at risk if you aimed to realize the ideal course of action instead. But the problem for White is that this last consideration is not ideally undefeated, as it also favors choosing the non-ideal course over the ideal one, due to your slightly inferior physical ability. That is, the reason you have for turning down the offer (*B*) is that if you were to accept it and aim to be courageous (*A*) then you'd likely end up realizing a course of action that was much worse (*C*), due to the risk that your cowardice would pose to public safety. However, because you are also less physically able than the next person in line for the job, the potential risk to public safety is *also* a reason to prefer the non-ideal course (*B*) over the ideal one (*A*)—which means it is ideally defeated.

It follows from this that, in (*h*), declining the offer on the grounds of your cowardice would be *opportunistic* according to White's definition, because considerations of public safety are ideally defeated reasons, as they favor the non-ideal course over both the disastrous *and* the ideal ones. (Again, it is important to remember that what is opportunistic is *not* always self-serving!) On White's account, then, declining the offer would mean treating your anticipated cowardice as a reason to “opportunistically” pursue public safety when you would normally have to prioritize your family's good. And this is entirely implausible.

We conclude that White's account, like Marušić's, is inadequate to address our puzzle. Both are inadequate, because both demand an asymmetry between the ways we can factor in our and others' chance of failure. Thus, both give rise to cases where we are screened off from making an honest assessment of our own reliability compared with the reliability of others.

Objection 2: Both Positions Incorrectly Classify Variants of Our Cases

Consider a variant on White's case of Pierre:

- (i) Leanne has a bad habit of spilling secrets, and in order to prove this to her Leanne's friend has told her a funny but harmless secret and

bet her \$20 that she cannot keep it to herself until the end of the night. Leanne takes the bet, but later on she has the opportunity to share the secret in a way that will improve her standing (or improve the spirits of another friend, etc.), though not to a degree that would be worth forfeiting \$20.

If Leanne were certain that she could keep quiet, it would not be worth spilling the secret for this small advantage. However, if Leanne believes that she is very likely to spill the secret at some point, then it may be wise to “take the win” and forfeit the chance of winning the bet. That is to say, in this circumstance it seems acceptable for Leanne to factor in the likelihood that she is going to give in and to decide that she might as well do so at an opportune time. And our account gives a simple explanation of why this is: it’s because the normative consideration that Leanne’s decision is responsive to is a *reason* that is grounded in the prospect of a material benefit, rather than a *logos* that is grounded in the trust of a friend.

By contrast, both Marušić and White are committed to treating Leanne’s reasoning as no more legitimate in this case than in the original. Marušić must say this, because Leanne is still settling the question of whether she will keep the secret by considering the evidence about whether she will follow through on the decision to keep it, rather than by practical reasoning. And White must say the same, because the consideration that favors the non-ideal choice of spilling the secret now over keeping quiet now and spilling it later on—namely, that even though this will cost Leanne \$20, at least it will have some benefit—is one that also favors this non-ideal choice over the ideal one of keeping the secret throughout the night.

A similar move is available with any of the other cases where hedging seems inappropriate. Indeed, we saw this already in the pairs of cases (a)–(c) and (b)–(d). If your reason for avoiding snacking is that you’re trying to cut down on calories, then it might make sense to eat a small snack now in order to keep yourself from eating a bigger one—though you’d have to be alive to the possibility that this is just a rationalization. By contrast, if your reason is that you are keeping a strict religious fast, then reasoning of this kind appears perverse. And this is part of what explains why *logos*-based reasoning is not a rare exception arising only in cases involving moral constraints, but a ubiquitous feature of our lives. If you’re trying to quit smoking and you frame this just as a matter of avoiding costs to your health that outweigh the pleasure of the occasional cigarette, it will be very easy to rationalize the not-so-occasional smoke. Given the seductiveness of such rationalizations, most of us are better served by Anscombean *logoi*, principles that simply preclude weighing the costs and benefits of doing

the things we think we should not do.³⁶ Sometimes the operative logoi are given by the moral law, sometimes by the state, sometimes by the inventor of a game, sometimes by one's community, and sometimes by the agent herself. Having given yourself a logos, it will preclude you from ordinary cases of hedging, at least so far as you remain committed to the rule being the best way to structure your deliberation.

4. Conclusion

In some cases, but *only* some cases, the choice of a suboptimal course of action cannot be justified by the undesired consequences that would ensue from a failure to follow through on the choice of something better. Our argument in this paper has been that what explains this phenomenon is the way that practical deliberation needs to be responsive, not only to the anticipated consequences of various possible choices, but also to practical principles that structure which courses of action are to be treated as live possibilities in the first place. There are certain actions—ones like violating the rules of a game, taking a person's property without their permission, or breaking a promise—that are simply not up for deliberative consideration, at least if they are done for certain reasons; and plausibly certain others—like killing an innocent human being—that are ruled out of consideration absolutely. These things are ruled out because they are *things that a person is not to do*, and not because they are things that one is obligated to *prevent* anyone else (or even himself or herself) from doing. Nothing in this implies that we should trust ourselves more than we trust others, nor that we should take account of others' defects of character any more or less than our own. Because of this, we are not only permitted but positively *required* to consider our own anticipated weakness of will in all the same ways we consider the weakness of others. This does not, however, generally license the violation of logos-type principles.

Of course, there is much more to say on this topic. There are two things, in particular, that we have not tried to do in this paper:

First, while we have made extensive use of Anscombe's distinction between reasons and logoi, and done our best to make clear how she understands it, we have not made any sustained argument for the distinction, nor have we shown why logoi are peculiarly connected with an agent's first-personal deliberative standpoint. That the reason/logos distinction proves so fruitful in resolving our puzzle suggests to us that the distinction carves the normative domain at its

36. On the moral and psychological importance of this kind of reasoning, see Richard Holton's discussion of what he calls *resolutions* in *Willing, Wanting, Waiting* (2009). For critical discussion of Holton's account see Bratman (2014) and Schwenkler (2024).

joints. Still, our solution to the puzzle is only as secure as its foundation, and it would take further work to establish the legitimacy of this distinction and respond to any objections to the arguments for it.

Second, nothing we have said addresses the very difficult question of how to reconcile the practical and theoretical perspectives on one and the same action. We have made a case that practical deliberation about what to do now (e.g., whether to refrain from eating a snack) can incorporate theoretical evidence about what one might likely do later on, conditional on what one now decides (e.g., whether skipping the snack now means that you'll eat a worse snack later). However, this is not enough to settle whether it is possible to integrate these perspectives, or to take them up simultaneously even when they are in conflict, in cases where they concern one and the same act—that is, in cases where a person has decided, or must decide whether, to do the *very* thing she has reason to believe that in the end she will not do.³⁷ As long as this remains unsettled, we will not have provided any general account of the integration of the practical and theoretical perspectives. Our suspicion is that there are profound antinomies which make it impossible for a knowingly weak-willed agent to have a fully integrated and rational perspective on her future actions in such a case. But this remains a topic for future work.

Acknowledgements

We are especially grateful to Robert Audi, Nir Ben Moshe, Luca Ferrero, Beri Marušić, and Maura Tumulty—as well as two anonymous referees at *Ergo*—for feedback on this paper. We also thank members of the Practical Philosophy Workshop at the University of Illinois, and participants at the 2023 St. Louis Annual Conference on Reasons and Rationality, for their helpful questions and suggestions. Finally, JS acknowledges research support from the Notre Dame Institute for Advanced Study.

37. Cases like this are at the center of Marušić's *Evidence and Agency*, and nothing we have argued here is in conflict with Marušić's contention that a full intention or sincere promise to do something is incompatible with the belief that one will not do it. What we have done is to challenge his argument that the demands of the deliberative standpoint preclude making certain decisions on the basis of evidence about what is likely to happen if one decides otherwise. Despite this, it might be that the person who has, e.g., decided to quit smoking or get a paper finished by the deadline, must believe that she will do this even in the face of countervailing evidence, if her decision is to be wholehearted. Likewise, it might be that the person who is morally bound to make a certain decision—as, e.g., a married person is bound to be committed to fidelity, and a firefighter on the scene of a dangerous fire must commit to acting courageously in order to perform a necessary rescue—must then believe that she will act in the obligatory way. For further development of this position, see Marušić & Schwenkler (2018; 2022) and Schwenkler (2019: 203–210); and for important background see Anscombe (2001: 91–94), Hampshire (2015), and Moran (2001).

References

- Anscombe, G. E. M. (1981a). Modern Moral Philosophy. In *Ethics, Religion and Politics: The Collected Philosophical Papers of G. E. M. Anscombe, Vol. 3* (26–42). Basil Blackwell.
- Anscombe, G. E. M. (1981b). On Promising and its Justice, and Whether it Need be Respected in *Foro Interno*. In *Ethics, Religion and Politics: The Collected Philosophical Papers of G. E. M. Anscombe, Vol. 3* (10–21). Basil Blackwell.
- Anscombe, G. E. M. (1981c). The Question of Linguistic Idealism. In *From Parmenides to Wittgenstein: The Collected Philosophical Papers of G. E. M. Anscombe, Vol. 1* (112–134). Basil Blackwell.
- Anscombe, G. E. M. (1981d). Rules, Rights, and Promises. In *Ethics, Religion and Politics: The Collected Philosophical Papers of G. E. M. Anscombe, Vol. 3* (97–106). Basil Blackwell.
- Anscombe, G. E. M. (1981e). On the Source of the Authority of the State. In *Ethics, Religion and Politics: The Collected Philosophical Papers of G. E. M. Anscombe, Vol. 3* (130–156). Basil Blackwell.
- Anscombe, G. E. M. (2001). *Intention* (2nd ed.). Harvard University Press.
- Anscombe, G. E. M. (2006). Knowledge and Reverence for Human Life. In Mary Geach and Luke Gormally (Eds.), *Human Life, Action and Ethics: Essays by G.E.M. Anscombe* (59–66). Imprint Academic.
- Bierson, Marshall (2024). Understanding Anscombe’s Absolutism. In Nathan Hauthaler and Nicholas Ogle (Eds.), *Anscombe and the Anscombe Archive* (97–120). Collegium Institute for Catholic Thought and Culture. <https://philpapers.org/archive/BIEUAA.pdf>
- Bierson, Marshall and Tucker Sigourney (2023). Famine, Affluence, and Aquinas. *Journal of Ethics and Social Philosophy*, 25(2), 307–322. <https://doi.org/10.26556/jesp.v25i2.1843>
- Bratman, Michael (2014). Temptation and the Agent’s Standpoint. *Inquiry*, 57(3), 293–310. <https://doi.org/10.1080/0020174x.2014.894271>
- Dworkin, Ronald (2001). *Taking Rights Seriously*. Harvard University Press.
- Faulconbridge, Peter (2021). Anscombe and the Rules of Ordinary Practices. In Adrian Haddock and Rachael Wiseman (Eds.), *The Anscombean Mind* (442–468). Routledge. <https://doi.org/10.4324/9780429198601-25>
- Ford, Anton (2011). Action and Generality. In Anton Ford, Jennifer Hornsby, and Frederick Stoutland (Eds.), *Essays on Anscombe’s Intention* (76–104). Harvard University Press. <https://doi.org/10.4159/harvard.9780674060913.c3>
- Gaita, Raimond (2004). *Good and Evil: An Absolute Conception*. Routledge.
- Goldman, Holly S. (1978). Doing the Best One Can. In Alvin I. Goldman and Jaegwon Kim (Eds.), *Values and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt* (185–214). D. Reidel Publishing Company. https://doi.org/10.1007/978-94-015-7634-5_11
- Hampshire, Stuart (2015). *Freedom of the Individual: Expanded Edition*. Princeton Legacy Library.
- Hart, Herbert L. A. (2012). *The Concept of Law* (3rd ed.). Oxford University Press. <https://doi.org/10.1093/he/9780199644704.001.0001>
- Holton, Richard (2009). *Willing, Wanting, Waiting*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199214570.001.0001>
- Huemer, Michael (2010). Lexical Priority and the Problem of Risk. *Pacific Philosophical Quarterly*, 91(3), 332–351. <https://doi.org/10.1111/j.1468-0114.2010.01370.x>

- Hyman, John (2015). *Action, Knowledge, and Will*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198735779.001.0001>
- Jackson, Frank and Robert Pargetter (1986). Oughts, Options, and Actualism. *The Philosophical Review*, 95(2), 233–255. <https://doi.org/10.2307/2185591>
- Johnson, Christa M. (2019). The Intrapersonal Paradox of Deontology. *Journal of Moral Philosophy*, 16(3), 279–301. <https://doi.org/10.1163/17455243-20182527>
- Kolodny, Niko and R. Jay Wallace (2003). Promises and Practices Revisited. *Philosophy & Public Affairs*, 31(2), 119–154. <https://doi.org/10.1111/j.1088-4963.2003.00119.x>
- Lazar, Seth (2015). Risky Killing and the Ethics of War. *Ethics*, 126(1), 91–117. <https://doi.org/10.1086/682191>
- Marušić, Berislav (2012). Belief and Difficult Action. *Philosophers' Imprint*, 12(18), 1–30. <http://hdl.handle.net/2027/spo.3521354.0012.018>
- Marušić, Berislav (2013). Promising Against the Evidence. *Ethics*, 123(2), 292–317. <https://doi.org/10.1086/668704>
- Marušić, Berislav (2015). *Evidence and Agency: Norms of Belief for Promising and Resolving*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198714040.001.0001>
- Marušić, Berislav and John Schwenkler (2018). Intending Is Believing: A Defense of Strong Cognitivism. *Analytic Philosophy*, 59(3), 309–340. <https://doi.org/10.1111/phib.12133>
- Marušić, Berislav and John Schwenkler (2022). Agency and Evidence. In Luca Ferrero (Ed.), *The Routledge Handbook of Philosophy of Agency* (244–252). Routledge. <https://doi.org/10.4324/9780429202131-28>
- McDowell, John (1979). Virtue and Reason. *The Monist*, 62(3), 331–350. <https://doi.org/10.5840/monist197962319>
- Moran, Richard (2001). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton University Press. <https://doi.org/10.1515/9781400842971>
- Rawls, John (1955). Two Concepts of Rules. *The Philosophical Review*, 64(1), 3–32. <https://doi.org/10.2307/2182230>
- Raz, Joseph (1999). *Practical Reason and Norms*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198268345.001.0001>
- Scanlon, Thomas M. (2000). *What We Owe to Each Other*. Harvard University Press. <https://doi.org/10.2307/j.ctv134vmrn>
- Schwenkler, John (2019). *Anscombe's Intention: A Guide*. Oxford University Press.
- Schwenkler, John (2022). Steering Clear of Trouble. *Philosophic Exchange*, 2021–2022. <https://soar.suny.edu/handle/20.500.12648/7676>
- Schwenkler, John (2024). How Temptation Works. *Journal of Ethics and Social Philosophy*, 27(3), 443–468. <https://doi.org/10.26556/jesp.v27i3.2920>
- Tarsney, Christian (2018). Moral Uncertainty for Deontologists. *Ethical Theory and Moral Practice*, 21(3), 505–520. <https://doi.org/10.1007/s10677-018-9924-4>
- Thompson, Michael (2008). *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press.
- Timmerman, Travis and Yishai Cohen (2020). Actualism and Possibilism in Ethics. In Edward Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2020/entries/actualism-possibilism-ethics>

- Vogler, Candace (2016). Anscombe on Promising. In Luke Gormally, David Albert Jones, and Roger Teichmann (Eds.), *The Moral Philosophy of Elizabeth Anscombe* (137–153). Imprint Academic.
- White, Stephen J. (2021). Self-prediction in Practical Reasoning: Its Role and Limits. *Noûs*, 55(4), 825–841. <https://doi.org/10.1111/nous.12333>