PREDICTION AND DECISION: LESSONS FROM THE HAPPY CASES

—draft—

Pamela Hieronymi

April 20, 2026

Much of the philosophical discussion about the interaction between theoretical and practical reasoning, or, as I will put it, between prediction and decision, tends to focus on what I will call the unhappy cases: cases of apparent irrationality, irresponsibility, or bad faith. One familiar character is Professor Procrastinate, who predicts that, due to his inveterate procrastination, he will not write the review he otherwise has strong reason to write. I would like to consider, instead, the happy cases: cases in which you predict you will do what you have strong reason to do. I hope thereby to better locate what has gone wrong in the unhappy ones.[1]

My ultimate quarry is the idea that prediction and decision, or theoretical and practical reasoning, are somehow isolated from one another, that they take place in different "standpoints," such that moving between them involves a shift in "point of view," and any attempt to combine them is somehow illicit or a sign of bad faith. The idea is inherited from Kant, who sought to render science and human freedom compatible by arguing that the activity of the rational will cannot be known by the methods of science. Kant was addressing the problem of free will by appeal to transcendental idealism and in-principle mystery. But the contemporary legacy of Kant's thought, in the metaphor of standpoints, leaves the idealism behind and extends to other contexts, where it brings more confusion than clarity. Or so I hope to suggest.

PREDICTION AND DECISION: REASONS AND QUESTIONS

We can start by distinguishing prediction and decision. You might predict that you will lose the match. This is different from deciding to throw the match. Both the prediction and the decision will leave you with what is, in some sense, the same view of your future: that you will lose.[2]

---

[1] While cases of breakdown can reveal important structure, they can also distort or hide what is otherwise in plain sight.

[2] The distinction is helpfully discussed in both G. E. M. Anscombe, *Intention* (Oxford: Blackwell Publishing Co., 1957). and Stuart Hampshire, *Freedom of the Individual* (New York: Harper and Row, 1965).

Some would say that when you predict, you adopt a "third-personal" "theoretical" stance or point of view on your future, while, when you decide, you occupy a "first-personal" or "practical" point of view towards your future. I believe we can do better by considering the reasons employed.

When you predict you will lose, you come to this view of your future by considering ordinary evidence—considerations that show it likely that your opponent will better you. In contrast, when you decide to lose, you do so by considering features of your situation that you take to count in favor of bringing about your own loss.[3] These reasons bear on different questions. The reasons you take to count in favor of acting are those you take to bear on the question of what *to* do. These are sometimes called practical reasons, bearing on the practical question. In contrast, the reason you take to show some outcome likely are those you take to bear on the predictive question, of what you *will* do. Such considerations are called evidence; they bear on what is sometimes called a theoretical question. The practical and the theoretical question are distinguished, not (merely) grammatically, but rather by the different sets of reasons that would bear on them—and so, I suggest, are prediction and decision.[4]

We can now say that there are, so to speak, two different routes to a view of your future loss.[5] You travel the first route by answering the ("theoretical") question of whether you *will* lose—where that is a question you could ask about anyone (whether you will lose, whether Luce will lose, whether Rodney will lose...). In settling this first question, you arrive at an ordinary belief, one which happens to be about yourself. The considerations you use to settle the first question (if you use any) will be those you take to show it *likely* that you (or Luce, or Rodney) will lose. You travel the

---

[3] More precisely, if you predict or decide for reasons, you do so for such reasons.

[4] This is the lesson I take from my attempts to explain why it is impossible (not just irrational) to believe at will—there is an irreducible difference in the reasons (which in turn makes for an irreducible difference in answerability), and the difference in the states of mind is to be marked by that difference in reasons (and answerability).

It may be worth noting that some considerations that one might take to bear on a theoretical questions are not well characterized as evidence, such as another person's say-so. I owe this point to Richard Moran. See Richard Moran, "Getting Told and Being Believed," *Philosopher's Imprint* 5, no. 5 (2005).

[5] I owe the metaphor to Hampshire, *Freedom of the Individual*.

second route by answering a different question—not whether you *will* lose, but, rather, whether *to* lose. This second question is not, so to speak, about anyone,[6] and so cannot be asked about anyone else. It is, in some sense, essentially "first-personal." In settling this second question, you arrive at an intention to lose. And whatever considerations you use to settle the second question will be—*in virtue* of your so using them—considerations you take to count in favor of (or against) losing.[7] Strikingly, the decision, no less than the prediction, also leaves you with a view of your future—but one grounded in your reasons for acting rather than evidence of likelihood.[8]

Notice, next, that prediction and decision remain stubbornly distinct. To see this, consider a radical case: Suppose that, seeing a flyer posted on campus, you volunteer for one of those experiments run by the psychology department, hoping for some free pizza.[9] A team of crack neuroscientists implant a chip in your brain by which they can control your thoughts and thus your behavior. They assure you that they will not make you think or do anything that would be highly unusual or out of character for you. They let you know that they plan to send you for a walk around campus at lunchtime, and you believe them. You thus predict that you will go walking.

---

[6] I am tempted to say, to answer the practical question is not to ascribe a predicate to a subject and so affirm a proposition, but rather to commit yourself to effect some represented change in the world.

[7] This is not to say that you *believe* they count in favor of or against *x*-ing. It is rather that, in *using* them to settle, positively, the question of whether to *x*, by employing them, you treat them as counting in favor of or against *x*-ing. Thanks to Eric Wiland for helpful questions.

It may be that some considerations that you use to settle the question of whether to φ do not *directly* count either in favor of or against φ-ing, by being costs or benefits, but rather undercut or enable other considerations. These still bear on the question of whether to φ, via those more complex relations.

[8] Traveling in reverse, prediction becomes (something like) explanation while decision becomes (something like) (something like) justification.

Some will think that my practical view of my future must, in some way, be based on evidence of likelihood, supported by my reliability in bringing about what I take myself to have reason to. I would argue, instead, that we have what Tyler Burge might call a default entitlement to rely on our own agency. Until shown wrong, we can move directly from practical reasons for φ-ing to a decision, which is to say, to a view of the future in which we φ.

[9] This is a slight variation on a case introduced to me by Christine Korsgaard.

But your prediction, alone, will not get you walking. If you are to go for a walk *intentionally*—if the neuroscientists are to get you to walk by controlling your mind, rather than just your body—then you will have to go for a walk because you *mean* to; you will have to decide to go for a walk. So, if the neuroscientists are going to make you walk intentionally, they also need to make you decide to walk. And predicting that you are going to make a decision is not the same as making it. More, when you make the decision to walk, you will not do so by appeal to evidence. You will do so by appeal to considerations that you take to count in favor of walking.

The fact that prediction and decision are distinct, that no amount of *knowing* amounts to *deciding*, has been thought to support the metaphor of standpoints. And his metaphor has, in turn, supported two further claims. The first is that the deliverances of the theoretical point of view can (or must) be ignored, when making one's decisions (I believe this is Kant's idea that we act "under the Idea of Freedom"). The second is that, when you make a decision about something, you must not then, or from within the same "point of view," regard it as certain. Thus, insofar as you regard some aspect of the future as certain, it is impossible, irrational, unreasonable, or at least somehow problematic to make a decision about it—at least from within the same "point of view."[10]

But both the metaphor and the further claims go too far. Prediction and decision routinely interact. Good decision-making often requires making predictions about yourself—whether you are likely to choke in the clutch or to forget your password. You might decide to throw the match because you predict you will lose it anyway, and you would like to save your strength. You might decide to go for a walk because you believe the neuroscientists will make you walk and you would rather not wait around any longer. Even though a prediction cannot simply stand in for a decision, it does not follow that a prediction must be ignored or lost sight of when making one's decision. Prediction, as such, is not isolated from decision-making, in some separate "point of view."

To see this, let us turn to the happy cases.

---

[10] The incompatibility of the two points of view is crucial to establishing "two standpoints" compatibilism.

HAPPY CERTAINTY

In the happy cases, you can predict with certainty that you will do something you also would have yourself to do. You are certain you will take the job or marry your partner—more, you are certain for the same reasons that your best friend or therapist is certain you will do so: you know what you care about and what you are like and you know what you will find convincing and why. Some of these cases provide, so to speak, the mirror image of the case of Professor Procrastinate: You are certain that your resolve will hold. In fact, your confidence in your future resolve may be a crucial part of your reason for deciding to embark down that path—to take the stand, be the whistle-blower, or adopt the child.[11]

I take the happy cases to be counter-examples to the claim that, insofar as you regard some aspect of the future as certain, it is then impossible, irrational, unreasonable, or at least somehow problematic to make a decision about it. My complete confidence that I will take the job or marry my partner is no bar to my decision to do so. In fact, it would be unreasonable not to make the decision.

An opponent might object: You might be certain that you will take the job or marry your partner, but your certainty cannot be the same kind of certainty enjoyed by your best friend—at least not when you occupy "your own point of view." Instead, your certainty, from your own point of view, will be the *result* of your decision—you can be certain about what you will do, because, as you think about your reasons, you have *already* made your decision. Your certainty is an instance of what is called "practical knowledge." More, in those cases in which your current decision depends on your confidence in your future decision, you are not *yet* certain of those future decisions. Rather, you are now certain only that, *if* you decide now to go down the path, then your resolve will hold. But you are not, prior to your decision, certain you *will* go down the path. So these happy cases are

[11] Note that, insofar as your prediction is, from your point of view, a crucial part of your reason for your decision, the prediction is not an instance of practical knowledge.

not, in fact, counterexamples to the claim that, if you regard some aspect of the future as certain, it is impossible, irrational, unreasonable, or at least somehow problematic to make a decision about it —or, that doing so requires, at least, shifting one's "point of view".

In arguing against the opponent's interpretation of the cases, I must grant three nearby truths. First: If you believe some aspect of your future is *impervious to your efforts*, then addressing the question of whether to (try to) change that aspect is (typically) unreasonable.[12] You should instead adopt what I will call the "fatalistic attitude" towards that aspect of your future: you should acknowledge it does not depend on your decision-making and plan around it. Our eventual death is like this. No matter what decisions we make, we will eventually die. It is thus unreasonable to make a decision about whether or not to die someday. We should, instead, take the fatalistic attitude towards our eventual death and plan around it.[13]

However, an aspect of the future can be *impervious to your efforts* without being *certain*, and the imperviousness, alone, is sufficient to render decision-making unreasonable. Perhaps you have been diagnosed with a genetic condition that carries a 30% chance of developing into a terminal illness, but none of your decisions will affect the chances of it so developing. The outcome is impervious to your efforts, though uncertain. That is sufficient to render a decision to (try to) change it as unreasonable. You should, again, adopt the fatalistic attitude.[14]

The happy cases are the inverse of the genetic condition: In the happy cases, you regard the outcome as *certain* but *not* as impervious to your decision-making. You believe the outcome depends

---

[12] "Typically" because of extreme cases. Perhaps there is no hope of swimming to shore. It may yet be reasonable to strike out. Reasonability is rarely absolutist.

[13] Of course, the *timing* of our death is not impervious to our efforts, and we can reasonably make decisions in an effort to delay it. And, I do not mean to rule out the possibility what one day humanity defeats death. But, insofar as I believe that such defeat is beyond the reach of my lifetime, I should take the fatalistic attitude.

[14] A further point to grant: Sometimes it is reasonable to "try" to do things that you believe you certainly will not succeed in doing—it is sometimes reasonable decide to do, or approximate, those actions that, in some more favorable conditions, or with more skill, might result in a different outcome—even when you are certain that, in current conditions, or given your current abilities, they will not. (This is only one use of "try.")

on your decision-making, and you regard it as certain in part because you regard your decision as certain. It would be entirely unreasonable, in these cases, to adopt the fatalistic attitude.

The second nearby truth to grant is this: Decision-making always involves the consideration of alternatives. If you are deciding whether to go for a walk, you are deciding whether *or not* to walk. A decision is the answering of a yes/no question, and such questions always admit of two answers: yes and no. So you must, in some sense, take there to be two possibilities before you: you could settle the question positively or negatively.[15]

This nearby truth might be thought to support the opponent's claim. It might seem that, when you address the question of whether or not to walk, you must entertain two possible, contrasting futures. And thus it might seem you must regard the future as open, awaiting your decision. But, it might be said, you cannot sensibly regard your future as both open and certain. That is some kind of contradiction. And thus it might seem that, insofar as you regard going for a walk as certain, you cannot sensibly address the question of whether to walk. (In fact, it will now follow that, if the neuroscientists are to get you walking intentionally, they will have to get you to give up on or at last "lose sight of" your firm prediction—at least momentarily.)

However, this argument relies on a falsehood and an equivocation. First, in order to settle the question of whether or not to act, you need not "entertain" the future in which you do not—at least in any robust sense of "entertain." When I settle the question of whether or not twice two is four, there is an alternative I close off: the negative answer. Yet I do not, in any robust sense, entertain the possibility that twice two is not four. (It is hard to know what it would be to entertain that possibility.) Likewise, when deciding whether to reassure my child or to tell the truth, I need not, in any robust sense, entertain the future in which I do not.

---

[15] Further, you understand that, if you settle the question of whether to walk positively, you should, and you usually will, work into the rest of your thinking and planning the fact that you will walk, while, if you settle it negatively, you should, and usually will, work into the rest of your thinking and planning the fact that you will not walk. Grant all this.

The falsehood leads to the equivocation: the sense in which you must regard the future as "open," when you make a decision, is only the one granted above: you must regard your future as depending (in part) on your decision—as not impervious to it. But it does not follow that it is unreasonable to regard your future as certain—because you might also regard your decision as certain. Thus the argument does not establish any contradiction in regarding your future as both *open*, in that it depends on your decision, and *certain*, because you are certain of your decision. Without showing such a contradiction, we do not yet have any reason to insist that we must shift to another "point of view."[16] (And so no reason to insist that the neuroscientists must prevent you from having confidence in their abilities in order to succeed.)

The third nearby truth to grant is this: My best friend predicts I will act because she knows how I will reason with the facts at hand. So, if I am making a prediction on the same grounds as my best friend, I am also considering the facts at hand. But those are the same facts we both believe I will find sufficient to decide. It seems, then, that if I were to try to make a prediction, on this basis, I would (also?) end up making the decision, and *therein* arrive at practical knowledge.[17]

Thus, to provide a clearer counter-example, we need to find a case in which my confident prediction of how I will act is not based on my current appreciation of the strength of the reasons for that decision. I believe we can find such cases by appeal to testimony. Suppose my best friend, or my therapist, is confident that I will adopt the child, or tell the truth when the time comes, or embark on the fight for justice. And suppose I trust them—I confidently believe, on the basis of

---

[16] The argument in a different form: When deciding whether or not *to* act, you are, precisely, settling the question of whether you *will* act. But whether you *will* act depends on the very decision you are in the midst of making. Thus, you cannot sensibly or reasonably be confident that you *will* act *while* you are making that decision. In reply: I agree that, when deciding whether *to* act, you are settling the question of whether you *will* act, but if you are confident that you *will* act *because* you are confident that you will decide to do so, then your confidence in predicting your decision is no bar to making it.

[17] "find sufficient to settle" is "employ in settling," not "believe to be sufficient"

their testimony, that I will.[18] "Yes, I know I will," I admit to them.[19] Still, the decision is complicated and momentous, and I have not yet made it: I do not yet see my practical reasons in the way that my friend confidently predicts I will come to see them, and so I have not yet employed those considerations in coming to a decision. Though I confidently predict, I have not yet decided.

If the opponent is correct, then (either such cases are impossible, or), in a case like this, in order to make my decision, I must suspend, or somehow step away from, my confident prediction.[20] But this seems unnecessary. More, it seems backwards. As I work my way to my decision, slowly coming to see the reasons in favor of adopting the child, it seems that my confidence in my friend's prediction could only be reinforced: they really do know me well.

For a more fanciful example, suppose I employ the neuroscientists to help me with my exercise habit. I have trouble maintaining my resolve to go for my morning walk, and they tell me they can help: for a small fee, they will ensure that I will decide to go for a walk every morning at 7:30. I pay the fee, confident they will succeed. I am therefore confident that, tomorrow, I will decide to go walking. The opponent might say that, having made these arrangements, I have already decided to walk each morning. Or perhaps the scientists are merely a placebo. But the following seems perfectly possible: Sometime between waking up and 7:30, I notice the cold and dreary weather, the warm coffee, and the possibility of walking later in the sunny afternoon, and I lose my resolve to walk this morning (maybe, in this window, the scientists are distracted, or maybe they are saving battery in their chip). I decide to stay on my comfortable couch. Even so, I maintain my confidence that, somehow, come 7:30, scientists will ensure that I change my mind and decide to walk.[21] And,

---

[18] One might think that, here, my reasons are not the same as my best friend's reasons—my friend does not predict based on their testimony! True enough. But my prediction may now be like my second-best-friend's reasons—my second-best-friend may rely on the testimony of my best friend.

[19] Perhaps I even do a little planning, based on the confident prediction I have thus inherited.

[20] Or I must enter a different "point of view" or "standpoint."

[21] I suspect that this paragraph describes something analogous to the prayer life of many people.

sure enough, as 7:30 rolls around, I think, "eh, fine, I'll go for a walk." The opponent must insist that, if I make this decision, then at some point before (or at least right at) 7:30, I must lose, or at least lose touch with, my confidence in the scientists' abilities. But neither seems either natural or necessary. Better, I think, to simply allow that a confident prediction can happily co-exist with a prediction in those cases in which what you predict you will do and what you would have yourself decide to do align.

Once we allow that a confident prediction about a decision need not interfere with making that decision, we can also allow that my confident prediction that I will tell the truth can be among my practical reasons for deciding to take the stand. But notice, it now seems that the trouble with Procrastinate is not *simply* that he allows a prediction, as such, to infect his practical reasoning.. If evidence of one's own *steadfastness* can unproblematically be part of one's reasoning about what to do in the happy case, then any problem with evidence of one's own *weakness*, in the unhappy case, does not lie simply in its functioning as evidence supporting a prediction about one's own future decisions.[22] The problem, I suggest, lies in the features of the case that make it an unhappy one.

THE UNHAPPY CASES

Let us turn then to the unhappy cases, in which what I predict I will do is not what I would have myself to do. I am in the dark story in which I predict the neuroscientists will make me decide to do something I now despise. Or perhaps the Oracle tells me the Fates have determined that I will kill my father. Or maybe I simply know my resolve will not hold. I think I should let my child cry, but I know I will not, or, like Procrastinate, I believe I should complete the review in a timely way, but I know I will not. What, exactly, is going wrong in these cases?

---

[22] I believe I may be here disagreeing with Berislav Marušić's solution to his "problem of difficult action." Marušić argues that, in one's own case, evidence of one's own weakness is not to be employed in making predictions about one's future but instead is to be taken up in one's practical reasoning, in one's planning, as evidence of the difficulty of what one plans to do. I agree that it should serve the latter role, but I think it may also serve the former. This may leave me owing an answer to Marušić's terrific problem. See Berislav Marušić, *Evidence and Agency: Norms of Belief for Promising and Resolving* (Oxford: Oxford University Press, 2015).

Many things are going wrong—manipulation, in some cases; weakness of will, in others—but they share this feature: The full set of reasons I take to be powerful would, if I were to draw from them the conclusions I take them to support, give me an inconsistent view of my future. The two routes to my future lead to different futures. Yet there can be only one future.[23]

What is to be done, in the face of such inconsistent reasons?[24]

It would typically be best, of course, to find some strategy to make it possible for you to do what you take yourself to have most reason to do. Disable the neuroscientists' device. Leave your partner to care for your crying child and exit the scene. With Ulysses, tie yourself to the mast. Or, make a side bet to raise the stakes, for yourself, for completing the review. By adopting a good strategy you thereby change your reasons for the prediction, while still achieving the outcome your practical reasons recommend.[25] Your view of the future is now unitary, and you achieve the desired outcome, albeit with a bit more time and effort.

---

[23] This is analogous to a case in which I take myself to have powerful reasons supporting inconsistent beliefs—that light is particle and that light is a wave, or that the butler is guilty and the butler is innocent. There is only one world. In the face of strong but conflicting reasons for belief, I can avoid contradictory double-vision by suspending judgment, remaining agnostic (or by having beliefs about likelihood). In the face of strong reasons for an inconsistent view of my future, I can also, for a time, delay decision-making—but that option will not be available forever. Time will move forward and some version of my future will be realized.

(In Marušić's cases of difficult action, a decision must be made: one must resolve or promise.)

[24] Marušić claims that I need not find myself in the contradiction, because I need not take the facts *others* take as evidence of future failure to be such evidence, and so I need not come to the same conclusion as others. This seems correct in some of the cases. But I may, still, find myself convinced by my bookie, so to speak, or by the Oracle. In that case, what to do?

NTS: you do not interpret these, from the start, as cases of *problematic agency* (cf. manipulation), but as cases of having *problematic reasons* (a result of one's situation, in the manipulation cases).

[25] It may be that the reasons you need to provide yourself, in order to change your prediction, are not as strong as the reasons your bookie would need, to change his. It may be that you are entitled to give yourself the benefit of the doubt, so to speak. Cp. "Morally Bad Beliefs" (in preparation). (You also change your *practical* reasons, in these cases—you adopt a more complex strategy. But you keep your intended *outcome* fixed.)

Notice that, even in these least unhappy of the unhappy cases, you are, so to speak, "planning around" your predicted poor decision—you are treating it as a fixed point.[26] You have not, exactly, taken the *fatalistic* attitude towards your future decision—you have not treated it as an outcome impervious to your choices. After all, it was *not* impervious—you found a strategy to avoid it altogether. Yet, in planning around it, you are treating it as, in a way, fixed. And that, by itself, indicates some remaining unhappiness.

What, exactly, is this remaining unhappiness? When all is well, you do not treat your future decisions as something you need to plan around; you treat them, instead, as decisions you can, now, optimize, so to speak. If you now believe it would be best to let your child cry, you can, now, count on yourself to let your child cry. There is no need to bother your partner. If you now think you have most reason to write the review, you plan to write the review when the time comes. There is no need to shore up the future with extra incentives. If instead you need to treat some possible future decision as one you need to avoid or shore up with extra incentives, you treat it as though it is not yours to optimize. It is as if you are counting on something or someone other than you—other than the mind now making the plans—to make that problematic future decision. There is a kind of disunity here—you cannot count on your future self, so to speak. I regard that disunity as a familiar defect of agency, an ordinary kind of weakness of will.[27] It calls for a good strategy.

---

[26] This "planning around" seems to me different than the resolution that Marušić suggests: taking the reasons that ground the prediction to be evidence of difficulty in making one's plans. To suggest that the grounds for prediction are instead evidence of difficulty is, it seems to me, to suggest that you, now, think it would be too much of a bother, or involve too much effort, to take the route that would go through the decision. But the trouble with that route (it seems to me) is not the difficulty, in *that* sense. It might be the *least* difficult route, all things considered, if you could only count on yourself to make the decision when the time comes. The trouble with that route is not the difficulty, *per se*, but rather your own unreliability—you have strong reasons to predict you will change your mind and do the wrong thing, when the time comes.

[27] Some may want to articulate this point using the metaphor of points of view—from the point of view of decision-making, your future decisions are up to you. But doing so generates confusion. It suggests that the trouble in these cases is *simply* that I regard certain future decisions of mine as certain or fixed. But, as shown by the happy cases, I can do this unproblematically. The trouble in these cases is not simply that I treat the future decision as fixed, but that I treat it as fixed *in a way that I now regard as poor*—the trouble is the predicted disunity in my view of what is to be done.

Unfortunately, though, good strategies are not always forthcoming. It is not clear how to avoid the Fates. It may be the best Oedipus can do is to live in the contradiction while trying to resist the prediction—to not go gentle, so to speak, into the prediction,[28] but rather to move forward with resolve, his (strong reasons for the) prediction notwithstanding. What else could he do? (Note, Oedipus' resolve never wavered— the Fates had to resort to deception.[29])

Sometimes the same recommendation is made to Procrastinate: Do not go gently! Resolve to complete the review, the (strong reasons for) the prediction notwithstanding! But whether simply resisting the prediction is a good recommendation, in any given case, depends, in part, on what is at stake and possible risks to others. In Procrastinate's case, simply doubling down on his resolve, raging against the prediction, seems irresponsible. It would be better for him to opt for a next-best outcome—simply declining the invitation.[30]

Notice that, if Procrastinate opts for this next-best outcome, he unifies his anticipated future by changing—by conceding to the prediction—his view of which outcome he has most reason to (aim to) realize.[31] However, he still avoids making the predicted poor decision. He works around it.

In what we might call the unhappiness of cases, you cannot find a way to avoid the predicted poor decision, and so you concede *both* the outcome and the decision, so to speak: You take

---

[28] Dylan Thomas, "Do Not Go Gentle into That Good Night," in *The Poems of Dylan Thomas* (New Directions, 1939). Thomas recommends resisting death and so avoiding the fatalistic attitude.

[29] Note, too, that, unlike the other unhappy cases, Oedipus' prediction was not based any doubt about his own resolve. Perhaps surprisingly, "do not go gentle" seems appropriate the case of manipulation, but not of weakness. Resistance can be recommended only to the strong.

[30] Stephen White imagines a similar but simpler case: "[S]uppose 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 don'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. Arguably, the right thing to do under these circumstances is to refrain from making the call." He suggests writing an email instead. Stephen J. White, "Self-Prediction in Practical Reasoning: Its Role and Limits," ed. Kyla Ebels-Duggan and Berislav Marusic, *Responsibility and the Demands of Morality: Collected Papers* (Oxford University Press, 2025), https://doi.org/10.1093/9780191997273.003.0010. 184.

White suggests that you email her, instead. Here, you are again planning around a predicted poor decision but giving up on achieving the outcome your practical reasons would otherwise recommend. You are opting for a next-best outcome.

[31] The prediction has entered his practical reasoning both as a constraint and as something avoidable.

something closer to the fatalistic attitude towards your own future decision, treating it as something you can neither change nor avoid—you go gently into the predicted decision. This response seems to many people problematic, a case of "bad faith."

But, again, a variety of cases meet this description, and they are not all problematic in the same way. The responses that Sartre calls "bad faith" are (only) two of these varieties. As I understand him, Sartre imagines, first, someone who tries to ignore the fact that no amount of prediction will amount to a decision, who pretends that the prediction simply settles the matter, without input from them, so to speak. They thus attempt to avoid responsibility. But, because prediction will not amount to a decision, the action, when taken, *will* be the result of their decision. There is no avoiding involvement. And we all know this. Thus, this pretense requires some self-deception, as Sartre noted.[32] Sartre also criticizes those who attempt the reverse—who try to treat a decision as if it has the independence of a prediction. His gambler tries to treat his decision as something that would settle the future in advance, without his own ongoing upkeep[33] In deference to Sartre, let us call these two kinds of cases "bad faith proper."

Consider, next, Pierre, introduced to us by Stephen White:

> 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. Perhaps telling what he knows in present company would improve Pierre's standing with the popular crowd. Or maybe it would just relieve him of the pressure of keeping it to himself. Since he has little hope of remaining a loyal friend for long, there's not much to be gained by forgoing these benefits of immediate betrayal. Thus reasons Pierre.[34]

---

[32] This person tries to take the fatalistic attitude towards their own decision—they try to treat it as impervious to their decision-making. But it is their decision-making and so cannot be impervious to it. Jean-Paul Sartre, *Being and Nothingness*, trans. Sarah Richmond (Oxon: Routledge, 2018), PAGES.

[33] I am grateful to Berislav Marušić for help in understanding Sartre's gambler. CITE JPS. CITE Marušić?

[34] White, "Self-Prediction in Practical Reasoning: Its Role and Limits," PAGE

This case is jarring, a clear case of acting irresponsibly.[35] But it does not seem to involve bad faith *proper:* Pierre makes no attempt to deny or avoid his own role in his decision-making.

One might resist. One might say that planning to make a poor decision, by itself, involves some bad faith proper—some misunderstanding about your own role in the outcome and some failure to take responsibility. I disagree, and, interestingly, it seems White does as well. In a footnote, he provides the following case:

> Pierre predicts that, if he doesn't reveal the secret now (to get it off his chest), he's liable to do so when it would do significantly more damage to his friend's reputation—or even to some third party's reputation. Because the value of preserving the other's reputation is in no way opposed to his doing the optimal thing and keeping silent, it's the kind of value that could provide a valid justification for doing the next best thing and betraying his friend's secret in a context where the reputational damage would be minimal. (200, fn 31)

Here, Pierre includes the poor decision as part of his next-best plan, but he does so in order to minimize the damage the poor decision will cause.[36] There is no suggestion that Pierre is seeking either to avoid responsibility or to deny his own role in the outcome. If this second Pierre is acceptable, then including a poor decision in a next-best plan is not always bad faith.

What, then, has gone wrong with the first, Problematic Pierre? White locates the problem, not simply in the fact that Problematic Pierre includes the poor decision in his next-best plan, but rather

---

[35] Note that White avoids saying that Pierre *promises* to keep the secret. Pierre should not promise! But the friend may not have known Pierre's untrustworthiness in this matter and may have divulged without securing a promise, on the assumption that Pierre, like most, would not share. Let this be the case.

White also provides a second case that does not involve another person: "Imagine that my doctor has told me that I need to keep a strict diet in order to bring my cholesterol down to a healthy level. But I've never been good at keeping diets and I think it quite unlikely I'll stick with this one long enough to do much good. So, I figure there's little point in passing up the delicious, extra-cheese pizza that's being served for lunch."

[36] This is similar to the case in which you decide to walk now because you believe the neuroscientists will make you walk and you would rather not wait any longer. It is unlike that case insofar as you do not regard walking as a poor decision.

Pierre does not take the fatalistic attitude towards his decision. He does not treat it as impervious to his decision-making.

I suspect that what makes this case seem acceptable is, at least in part, the fact that it does not suggest that Pierre is giving into a moment of temptation. It suggests prior planning. In contrast, the other two cases strongly suggest giving into temptation, at the moment, with a rationalization—that is, with a line of reasoning that you would not have accepted prior to the temptation.

in the reasons for which he makes that next-best plan. White argues that, in planning our next-best action, we are subject to what he calls an "anti-opportunism" constraint. He says,

> The intuition behind this proposal is that we should rule out forms of reasoning that would enable a person to take advantage of anticipated wrongdoing or imprudence to secure benefits she would have to forgo if she were more virtuous, rational, or strong-willed…. [Pierre's] interest in getting a good laugh… 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… 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] 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 to take toward this flaw in his character. (200)

The problem with Problematic Pierre, as White sees it, is that he treats his own weakness as allowing him an opportunity to indulge his temptation. I would put it slightly differently, saying that your reasoning to your next-best case should remain constrained by the reasons (or values) you would have yourself act on, if you could count on yourself.[37]

So far we have considered cases in which you find ways to *avoid* your predicted poor decision, perhaps while still securing the most desirable outcome or perhaps while opting for a next-best outcome; cases in which you live in the contradiction while raging against it; cases of bad-faith proper, in which you try to pretend that there is only one route to your future, so to speak—either by pretending the prediction takes the place of the decision or that the decision has the independence of a prediction; and, finally, cases in which you work a predicted poor decision into some next-best outcome. If you do the latter, your reasoning to the next best outcome should not rely on the considerations that you take to tell against the poor decision.

But we have yet to consider what one might have thought to be the hardest and unhappiest case: you do not adopt any next-best plan; you have not found a way to minimize the damage; the Fates

---

[37] It seems there are two separable ideas here. White characterizes Pierre as making use of his own weakness as an excuse, so to speak, which is a "problematic attitude to take towards a flaw in his character." My alternative does not turn on an attitude towards your character flaw. The case of Unproblematic Pierre seems to shows that reasoning about one's own weakness and taking it to be a reason to act (an "excuse" for acting?) is not *itself* problematic. I would locate the problem in the reasons Problematic Pierre takes himself to have: He thinks that a future decision would be poor, but he plans to avoid that decisions by acting on the kinds of reasons that would motivate it—reasons that, even now, he thinks would be poor reasons for the future decision.

have not deceived you; the time for decision clearly arises; and so, here you are, now facing the temptation you predicted you will not resist. What now?

If you do anything intentionally, you will make a decision.[38] If the prediction turns out to have been correct, you will do so poorly. The reasons for action that you believed to be most powerful will not be those you now employ, in deciding to act. Either you will, at least temporarily, change your beliefs about your reasons, or else you will simply choose to do what you also believe you have to have most reason *not* to do. In the latter case, the reasons you employ, in deciding, will not be the reasons you recognize, in believing.[39] That is to say, the considerations you recognize, when you reflectively form a belief about what you have most reason to do, are not the considerations you then employ, when you decide whether to do it. While you *ought* to employ, in your decision-making, the same considerations you recognize, in believing, you need not: it is possible (though not rational), to be disunified in this way—to believe certain considerations provide strong reasons for acting without employing those reasons in deciding to act. Such disunity is a defect. It is not best characterized as adopting a different "standpoint" or "point of view." So long as you are not also engaging in the pretenses and self-deception of bad faith proper, these are simply ordinary cases of weakness of will. They are, I would claim, no more puzzling for having been entirely predictable.

CONCLUSION

So, where are we left? I hope I have shown that clarity is gained, and unclarity lost, by retiring the metaphor "points of view" or "standpoints" and replacing it with an appeal to different questions and different sets of reasons that bear on them. By distinguishing between the theoretical question, of whether you will act, and the practical question, of whether to act, we can both explain the

---

[38] Simply predicting will not amount to acting. To not make a decision is to decide. And, if you decide the basis of reasons, you will do so on the basis of considerations you take to count in favor of acting. If you make a decision, any reasons for which you do so will, in virtue of their use in your decision, be ones you took to count in favor of acting. NTS: here you might be interpreted into your decision, so to speak, in your attempted passivity.

[39] I give this interpretation of cases of weakness of will in Pamela Hieronymi, "The Will as Reason," *Philosophical Perspectives* 23 (2009).

stubborn distinctness of prediction and decision while also allowing them to interact in the ways they readily seem to interact. We can then see that predicting our own decisions, and even incorporating our own predictions about our future decisions into our plans, is not *in itself* problematic. Rather, problems arise only when what we predict we will decide and what we would have ourselves decide do fail to align. And, by resisting the over-broad visual metaphors, we also avoid painting all the unhappy cases with the same brush. Although all such cases display some sort of disunity, and so some defect of agency, not every case in which one concedes to a prediction is an attempt to evade responsibility. Some cases are doing the best one can, given one's defects, acknowledging and taking responsibility for them.[40]

[40] Material here has long gestated (under different titles) and owes many debts, many of which I have surely forgotten. I received helpful feedback from audiences at Princeton, Berkeley, the University Toronto, NUSTEP, SLACCR, the Eastern APA in 2016, and the Central APA in 2026. I received extremely helpful written comments from Maura Tumulty, Eric Wiland, and Stephen White. Berislav Marušić and I have been discussing these topics for years. ADD.

BIBLIOGRAPHY

Anscombe, G. E. M. *Intention.* Oxford: Blackwell Publishing Co., 1957.

Hampshire, Stuart. *Freedom of the Individual.* New York: Harper and Row, 1965.

Hieronymi, Pamela. "The Will as Reason." *Philosophical Perspectives* 23 (2009): 201–20.

Marušić, Berislav. *Evidence and Agency: Norms of Belief for Promising and Resolving.* Oxford: Oxford University Press, 2015.

Moran, Richard. "Getting Told and Being Believed." *Philosopher's Imprint* 5, no. 5 (August 2005): 1–28.

Sartre, Jean-Paul. *Being and Nothingness.* Translated by Sarah Richmond. Oxon: Routledge, 2018. 1943.

Thomas, Dylan. "Do Not Go Gentle into That Good Night." In *The Poems of Dylan Thomas*: New Directions, 1939.

White, Stephen J. "Self-Prediction in Practical Reasoning: Its Role and Limits." In *Responsibility and the Demands of Morality: Collected Papers,* edited by Kyla Ebels-Duggan and Berislav MarusicOxford University Press, 2025. https://doi.org/10.1093/9780191997273.003.0010.