Deliberation and prediction: it's complicated.

Katia Vavova
Mount Holyoke College

Some think that deliberation crowds out prediction—that when you are deliberating about what to do, it is not rational to take into account evidence or predictions about what you are likely to do. Alan Hájek calls this the DARC thesis, and understands it as saying that Deliberation Annihilates Rational Credences—specifically option-credences, or credences about your options. The DARC thesis, he argues, is an implausible and theoretically bankrupt view: it conflicts with mainstream commitments in epistemology and decision theory, has absurd consequences, and only bad arguments to commend it. After masterfully attacking argument after argument for it, Hájek concludes that the DARC thesis is false and that there is no problem with deliberation and prediction. Deliberation, Hájek happily concludes, doesn’t crowd out or annihilate prediction—it welcomes it.

Hájek’s attacks on the arguments he considers for DARC are formidable. He thus undermines much of the available support for the thesis. He does not, however, show that there is no problem for deliberation and prediction, and he hardly establishes that deliberation welcomes prediction. But this is not the only way in which Hájek’s treatment of the DARC thesis falls short.

First, Hájek doesn’t consider the most plausible version of the thesis. Second, he gives a brief and superficial treatment of the best argument for it: the, what he calls, “wishy-washy agency/free will argument”, which he parodies rather than reconstructing. Finally, Hájek fails to consider motivations for something like the DARC thesis—why anyone would ever believe anything like it in the first place. In doing so, he misses a deep puzzle about deliberation and prediction—a puzzle which all of us, as agents, face, and which we may be able to resolve by accepting something like DARC. Thus, even if Hájek’s attacks on the arguments for DARC are totally destructive, we are left amongst the rubble not knowing what to think.

Maybe that’s not Hájek’s problem—it’s impressive enough to accomplish utter destruction, even without a further, positive proposal. Nevertheless, if we are going to reject a view, we should be clear about the consequences. And recognizing the consequences, in this case, should make us rethink our rejection and rethink the view. Maybe there is a better, more plausible version of it that doesn’t leave us in such a lurch.

My goal here is thus twofold: to show that there is a puzzle about deliberation and prediction, which would be solved by severing the link between them, and to sketch a way

---

1 See Hájek’s “Deliberation Welcomes Prediction” in this volume.
2 I take this to be the version defended in Marušić 2015, which I closely follow in this piece.
3 Hájek himself notes that he is parodying (see his section 5.1).
of severing the link that isn’t as absurd as Hájek thinks. In accomplishing these goals, I hope to both complicate—for the better—the debate between Hájek and his opponents, and to show that while deliberation might not annihilate prediction, it doesn’t welcome it either. Their relationship is, at best, complicated.

Here’s how I’ll proceed. I’ll start by considering a case of deliberation, which is importantly different from the sort that Hájek considers (such as which wine to order or what to eat for dinner). The difficulties about deliberation and prediction only come out when we look at cases of difficult action, e.g., ones in which we resolve to be faithful, acquire a healthy new habit, drop an unhealthy old one, et cetera. With these difficulties in our sight, I will consider some variant cases, which show that deliberation complicates prediction—even if they don’t show exactly how it does this. Recognizing this difference points to a better, more plausible version of the agency/free will argument. It also motivates us to accept something like DARC. I close with some thoughts about remaining difficulties.

1. Motivation; or, why anyone would ever be inclined to accept anything like the DARC thesis in the first place.

Consider the following case from Marušić (2015: 15).

**Marathon.** I’m thinking about running a marathon. But running a marathon is hard—and not just during the marathon. It requires serious dedication and training in the months before the race. This kind of training may be no big deal for otherwise athletic folks, with existing exercise or training habits. It may involve just a marginal increase in the time spent exercising, or perhaps just a new exercise routine. But I am not one of those otherwise athletic folks. I have maintained past exercise resolutions about 30% of the time. So, I can reasonably conclude that there’s a 30% chance that I won’t wimp out on this one.

In light of this fact about my likelihood of failure, we should ask two questions. First, can I rationally believe that I won’t wimp out on my marathon training? Second, should I resolve to run the marathon?

To believe that I won’t wimp out would be to believe against my evidence. But that seems like paradigm irrationality. So, I shouldn’t believe that I will follow through on my training. It would be irrational to do so. This answers our first question in the negative.

But if I believe that there’s a significant chance I won’t do something, then how can I sincerely resolve to do it? So, perhaps, if I believe that I will most likely wimp out of my marathon training, then I shouldn’t—or maybe even can’t—resolve to run the marathon. This answers our second question in the negative.

If Hájek is right, and deliberation really welcomes prediction, then we are in trouble. For
we make such resolutions all the time—whether at the altar when we promise to spend
the rest of our lives with one another, on New Year’s Eve when we resolve to be better
than before, or when we decide to quit smoking or drinking. More generally, we make
such resolutions whenever we decide to do something difficult. This is why such
resolutions are so important to us, both personally and socially. They make up the fabric
of our lives—lives that would be seriously impoverished, if not impossible, without them.

This, then, is our puzzle. We are doomed either to irrationality, insincerity, or to never
making decisions about such difficult matters. The first options are implausible and
unattractive; the third is unthinkable. We need a solution.

1.2 Urgency

The urgency of the puzzle is even clearer in the context of the problem of women and
underrepresented groups in philosophy. Two questions put the issue into focus. First,
how should individuals from such groups view their chances of success in the philosophy
world, given the statistics? Second, how should we mentor and advise such individuals?

To see the difficulty, start with the following case.

Philip. I am a statistical anomaly. As a philosopher and a woman, I am more
likely than male peers to quit philosophy after taking Intro, to choose against
graduate school, to drop out of graduate school if I do choose it, to fail to get a
tenure-track job, etc. I am also more likely to choose family over career at any of
these points.

The problem for women in philosophy is starkest at the faculty level, where it is estimated
that only 16.6% of philosophers in full-time jobs are female. The numbers are worse for
some other underrepresented groups.4

The statistics support the assumptions in Philosopher, though they don’t, of course,
explain why women drop out at any of these points or show that their dropping out is
either bad or involuntary. However, the fact that a decision can be voluntary, rational,
and even good for the individual in the particular context doesn’t show that there isn’t a
problem. In unjust social structures, rational, self-preserving decisions by individuals may
still be problematic. And as Sally Haslanger puts it, “With these numbers, you don’t need
sexual harassment or racial harassment to prevent women and minorities from
succeeding, for alienation, loneliness, implicit bias, stereotype threat, microaggression,
and outright discrimination will do the job” (2013).

4 See the Digest of Education Statistics reports from 2003 (the most recent data compiled for philosophy).
See also Beebee and Saul 2011, Goddard et al. 2008, and Van Camp 2014. For a summary of some of the
problems, as well as hypotheses about
In light of these numbers and these facts, we can ask again, with respect to Philosopher, (1) whether I can rationally believe that I will succeed, and (2) whether I should resolve to try, given my likelihood of failure.\(^5\)

As in the marathon case, the first answer seems to be no: believing I’ll succeed requires believing against my evidence. Again, the second answer also seems to be no. If I’m not irrationally ignoring evidence, then I know that I’m likely to fail. But it’s hard, if not impossible, to resolve to do something that you think you won’t do.

These answers have, to put it mildly, unattractive implications for individual deliberations, like those in Philosopher, but also for our mentoring of women students and students of color. The options seem to be to either give up now and save ourselves the trouble or to plunge ahead blindly and irrationally believing that we will make it. (That belief may be required for us to actually have a chance of making it.) Are these really the best options for me and for my students, similarly haunted by stereotypes and statistics? And does this mean that we shouldn’t take measures to diversify syllabi, invite speakers from underrepresented groups, and the like, because this may fill our students with false hope, inviting them to love a discipline that doesn’t care about their voices? But then what should I do? Should I tell my students to give up and save themselves the trouble? Or, should I lie about their chance of success? Or, should I suggest they ignore the evidence and believe, irrationally, that they’ll probably succeed?

None of these are good options. And we won’t be able to find a better one without an account of the relevance of such statistical information—an account of the relationship between deliberation and prediction that allows the former to complicate the latter in the way that something like DARC suggests.

To reiterate the puzzle: if Hájek is right that DARC is false, then philosophers from underrepresented groups are doomed to either irrationality or giving up. If not unthinkable, this is at least implausible and depressing. We really need a solution.

Here’s an option: break the connection between deliberation and prediction. If we could show that these predictive facts about our likelihood of failure are somehow irrelevant to our deliberations about what to do, or have a different significance in the context of deliberation, then we might be able to resolve this puzzle. It is important, then, that we take DARC, or something like it, seriously, even if it has unattractive theoretical consequences and mostly bad arguments to recommend it. This puzzle suggests that there may be better arguments yet, and that the theoretical costs might be worth it.

2. Evidence that deliberation complicates prediction.

\(^5\) I mean ’failure’ here in a thin sense, where women who drop out of philosophy needn’t be seen as failures and where I can fail to go to the grocery store today simply by choosing not too.
We now have a motivation for accepting something like the DARC thesis. Cases like Marathon and Philosopher in which we deliberate about difficult matters push us into a puzzle that accepting something like DARC may help us resolve. But why should we think that deliberation actually gets in the way of prediction? Marathon and Philosopher show us it’d be nice if it did, but they don’t show us it does. To warm us up to the idea that deliberation complicates prediction consider a variant on Marathon.

**Lottery Marathon.** I’m trying to decide whether run a marathon. Whether I run it depends on whether I get in via the lottery. The chance of winning the lottery is 30%.\(^6\)

In the original marathon case, there was a 30% chance that I would run the marathon because that was the chance of my sticking with my resolution and not wimping out of training. In this case, there’s a 30% chance that I will run the marathon because that is the chance of my winning the marathon lottery.\(^7\) The chances of success and failure are thus the same in both cases, but for different reasons.

Two observations about this. First, it is clear that I can’t decide to win the lottery. It’s not just that the odds are low—it’s that it’s not up to me whether I win. Second, it is not similarly clear that I can’t decide to not wimp out. The odds are, again, low—equally low, but not wimping out seems like the sort of thing that is up to me. We often make such decisions, and sometimes we succeed—even when the odds are against us. Furthermore, these decisions are often sincere. Finally, believing that we will follow through doesn’t always seem irrational.

Thus, it seems that I should view my chance of failure differently in Marathon and Lottery Marathon, although the numbers are the same. The difference that makes a difference here is just the aforementioned: whether I fail to run is up to me in Marathon, but not Lottery Marathon. It is up to me whether I wimp out, but not up to me whether I win the lottery.\(^8\)

Notice that a well-informed observer whose only aim is prediction, say a bookie who “offers bets on the athletic achievements of philosophers” (Marušić 2015: 20), doesn’t have the same pressure to respect this difference between that which is up to me and that which isn’t. That is because this bookie’s only interest is in predicting what will happen, and that is just a numbers game. The bookie should, then, make the same bet in Marathon and Lottery Marathon.

---


\(^7\) I’m assuming that in Lottery Marathon, I am not a wimp—I have a perfect track record with exercise resolutions, so losing the lottery is the salient way I might fail to run.

\(^8\) We can construct a similar version of Philosopher where my chances of staying in philosophy depend not on my own efforts and dedication, but on whether I win the Department Will Support You Lottery. Even if my chances of staying in philosophy are equally low in Philosopher and Lottery Philosopher, I should view those chances in the same way because things are only up to me in the original case.
As the deliberating agent, however, I cannot similarly ignore the difference between these cases. If I were to ignore this difference in the course of my deliberations and bet against myself in both I would be making a mistake: I would be treating something that is up to me as if it isn’t. This is a factual mistake: though it may be hard for me to not wimp out, not wimping out is, in fact, up to me. This is also a practical mistake: involves deliberating about and acting toward something under my control as if it isn’t. It is, finally, a mistake with potentially serious consequences: in treating something that is under my control as if it isn’t, I may well ensure my own failure. Maybe this isn’t a big deal when it comes to whether I run a marathon—but it is a big deal with respect to whether I pursue and succeed in the career I hope to have.

This all suggests that there is a real and important difference between deliberation and prediction. Some ways of failing to act as we have resolved to involve failing to do what is up to us—because we got tired, wimped out, or simply changed our minds. Other ways of failing to do what we have resolved to do are not our failures—they are external impediments, such as losing a lottery or breaking a leg. We have now seen how recognizing this mundane difference helps motivate the thought that the same chance of failure can have very different significance, depending on whether we are agents or onlookers and, if we are agents, whether things are up to us or not.

Perhaps this much is right: deliberating about what to do affects the way we should view predictive evidence about our own future actions. But can’t we, as agents, also engage in the business of pure prediction about our own actions? If, for example, I go to the bookie to place bets on whether I will run in Marathon and Lottery, shouldn’t I bet just as the bookie does? After all, I am, like the bookie, just playing the numbers game. But if I can do this, then deliberation definitely doesn’t block or annihilate prediction. And if that’s right, then why should we continue to think that there is some important difference, even from the agent’s perspective, between deliberation and prediction?

This much is right: it is possible for agents to view their own actions as they would the actions of strangers, and to make bets accordingly. Thus, our situation as deliberating agents doesn’t make pure prediction impossible. But that was never the right way to understand the DARC thesis. The point isn’t that we can’t take this perspective on ourselves, but that we shouldn’t, because in doing so, we make a mistake. The mistake is that of treating something that is up to us as if it isn’t. In making this mistake, we ignore a relevant difference between the cases. We also treat ourselves as if we were strangers. This requires a kind of alienation from ourselves, as well as an implicit denial of our own agency.

Hájek worries, however, that this move overgeneralizes. The bookie, he says, clearly

---

9 Here we can see the sort of appeal to agency and freedom, which lies behind the so-called wishy-washy agency/free will arguments. I haven’t worked out such an argument here, but have rooted the appeal in real, concrete differences between cases. For one worked out version see Marušić 2015.

10 See Hájek’s sections 5.1 and 5.4.
doesn’t disrespect my agency! So why should we think I disrespect my agency in taking the same perspective on myself that the bookie does? But we should deny that the bookie doesn’t disrespect my agency. She does, and in just the same way that I do when I take on the purely predictive project with respect to myself. However, the bookie’s disrespect of my agency isn’t problematic as my own is. This isn’t because the bookie isn’t me. Rather, it is because she isn’t engaged in the practical question of what to do, but the theoretical question of what will happen. My status as agent may be irrelevant to the latter question, but it is relevant to the former. Similarly, then, if I am only interested in the theoretical question of what I am likely to do, I could also ignore my own status as an agent. That is, if I’m not engaged in deliberation, these likelihoods can play a different role for me. But we should be aware of the cost and mistake of taking this perspective on ourselves.

3. A better way to understand the relationship between deliberation and prediction.

The foregoing suggests that facts about my likelihood of success play a different role in different kinds of inquiry. In particular, they play a different role for me when I am engaged in deliberation than they would if I were engaged in mere prediction, like a disinterested well-informed observer. This provides a way of filling in the agency arguments and vindicating claims like the one that predicting what one will do disrespects one agency, or presupposes one’s non-agency, or something like that.

On the version of this view that I am sketching, deliberation doesn’t annihilate prediction, so it isn’t DARC. Still, deliberation complicates prediction (DCP). On this view, credences about my chance of success or failure aren’t relevant to my thinking in the same way that they would be if I weren’t deliberating.

Crucially, this doesn’t mean that such credences are wholly irrelevant to my thinking. On the contrary, statistical facts about the likelihood of failure are important for my deliberations. If I know that something is going to be difficult to do, that is something I

---

11 Marušić [2015] holds that taking the predictive stance toward others only disrespects their agency if we are close to them or have a significant relationship with them. So, I’m not disrespecting a stranger’s agency when I predict her behavior, but I am disrespecting my friend’s agency when I predict hers. I see no reason to constrain the view in this way, however. If part our mistake in taking the purely predictive stance toward ourselves is that of denying our agency, then anyone who takes the purely predictive stance toward us is also denying our agency and, therefore, also making a mistake—regardless of their relationship to us.

12 Hájek similarly worries that this move overgeneralizes to one’s future self. This is supposed to be a problem because relative to my future self, I am like a well-informed onlooker. So long as I’m not deliberating about what to do in the future, I am in a position to rationally predict what I will do—I am thus in a position to have option-credences. There doesn’t seem to be anything wrong with this, Hájek argues. But there is, and for the same reason as in the onlooker case: if I just focus on prediction I am denying, or ignoring my agency. Is that okay to do? It depends on what my project is. If my project is deliberating about what my future self will do, then yes, it’s a problem. If my project is mere prediction, then maybe not. (Perhaps, as agents, our project should never be mere prediction. But that requires further argument. Cf. Marušić [2015] section 6.1.)

13 See Louise [2009] and Marušić [2015].
can take into account when deciding if and how to do it. And statistical or track record
data about my past successes and failures, or those of people like me in similar situations,
can be relevant to this.

Consider an example. It is notoriously difficult for longtime smokers to quit. This is
important to recognize if you are a longtime smoker who wants to quit. Ignoring it might
lead you to try to quit in ways that guarantee failure—by going cold turkey with no
support system in place or by quitting during an already difficult and stressful time of
your life. Or, consider the Philosopher case again, in which there isn’t a complicating
factor of addiction. There, knowing the statistics about and complications of being a
woman in philosophy might help me approach my project in the right way: with, perhaps
again, a support system to help me get through the rough patches.

Thus, evidence that your chance of failure is high is important for deliberation. If that’s
right, then deliberation doesn’t crowd out or annihilate prediction. It’s not that it is
impossible to have credences about what is likely to happen when I’m deliberating. It’s
also not that such credences are irrational. They just aren’t relevant to my thinking in the
same way that they would be if I weren’t deliberating. Crucially, the fact that these
judgments aren’t relevant in one respect does not show that they are irrelevant. Just as it
is a mistake for me to treat things that are up to me as if they are not, it is a mistake to
treat things that are going to be difficult as if they are going to be easy.

We must thus somehow respect the statistical evidence about our chance of failure. But
how are we to do this without succumbing to the technical and theoretical problems
Hájek points out? Consider one such problem: probability gaps. You have a probability
gap if there’s something about which you don’t have a probability. Hájek argues that
DARC entails probability gaps because it entails that you can’t have rational credences in
propositions about your options—about what you will do. In other words, I can’t believe,
in Marathon, that I have a 30% chance of failure. There is nothing where that credence
would be.

Hájek argues that probability gaps are a problem for a number of pillars of Bayesian
epistemology including conditionalization, conditional probability, decision theory, and
others. He is right. These various formulas output counterintuitive results if our
probabilities are undefined. This is one of Hájek’s primary reasons for rejecting DARC: it
has bizarre and theoretically disastrous consequences for our epistemology.

If deliberation doesn’t annihilate prediction, however, but merely complicate it, it isn’t
clear that we are saddled with such gaps. Although Lottery Marathon shows us that
statistical evidence must be treated differently in deliberation, it doesn’t show us that such
evidence is irrelevant. Furthermore, the importance of such evidence for good
deliberation—and for not making the other mistake of treating something that is difficult

---

14 In allowing option-credences themselves to play a role in deliberation in this way, I part ways with
Marušić [2015] and present a view closer to his [2013].
as if it is easy—shows that the statistical evidence can’t be wholly irrelevant. If I am required to take this evidence into account in my deliberations, then I must have some attitude toward it.

It’s not clear exactly what that attitude would be, and how we would work this out formally. But let me sketch two options.

Recall, first, that my chance of failure is 30% in both Marathon and Lottery Marathon. So far, I’ve suggested that this number must have a different significance in the two cases because of that crucial difference: that my failure in only one of the cases is up to me. The mystery here is what exactly this “different significance is”.

One option is to say that although I believe that the objective chance of my failure in both cases is 30% my credence that I will fail is lower in one of the cases than in the other. Maybe the evidence that there is a 30% chance of failure plus something about my deliberative situation—perhaps the evidence that this is something that is up to me—entails that I should believe that my chance of failure is less than 30%. We might, then, explain the difference in my confidence in Marathon and Lottery Marathon as justified by the fact that free agents just aren’t constrained by statistics that predict their decisions as they are by statistics based on matters that externally constrain their decisions. This would violate the Principal Principle, so it has its own theoretical costs, but it isn’t crazy and it doesn’t obviously entail a probability gap.

Another option is to think there is a more complicated picture involving higher-order attitudes. For example, the statistical evidence in Marathon and Lottery are the same and they are evidence of a 30% chance of failure. However, the evidence in Marathon is weaker than the evidence in Lottery—so I should be less sure that this estimate of my failure is correct. Why? Perhaps it is, again, because I am an agent and running the marathon is, in the former case, up to me. This is a mere sketch, but the point for now is that it is worth exploring. And if something like it is right, then again there are no probability gaps, and we avoid some of Hájek’s disastrous consequences.

4. Implications for disagreement with Hájek.

I have sketched a subtler and, I think, more plausible version of Hájek’s target. Much more needs to be said to work out this way of understanding the idea that deliberation complicates prediction. For now, just notice a few upshots.

First, this view, DCP, doesn’t apply equally well to all of the cases Hájek discusses. Some of his cases don’t involve deliberation (e.g., the coin flip case in section 5.5). Others are too easy to be interesting or troubling (e.g., which wine to have with dinner in section 3). In easy cases like the latter, it is both up to you what to do and easy to convert decision into action. These cases are crucially unlike the difficult cases about, e.g., running marathons, promising lifelong fidelity, or quitting smoking. It is hard cases like these that
motivate the DCP thesis.

We should thus conclude that deliberation complicates prediction to the extent that the thing we’re deliberating about is hard to do. There’s a principled reason for this: the hard cases involve stuff that’s hard to do. We know it’s hard to do because we have option-credences that suggest that the likelihood of success is low. It is those cases that we need to know how to negotiate, and those cases that DCP is most relevant to.

If the above is right, then although deliberation complicates prediction, it doesn’t annihilate it. This brings us to our second point: DCP doesn’t entail probability gaps—at least not obviously. Facts about the likelihood of failure are relevant to our deliberations: they are what give us reason to think that what we are trying to do is difficult. And that is something I need to take into account when deliberating about what to do and how to do it. So I must have attitudes about them.

If DCP doesn’t entail probability gaps, then it doesn’t have many of the bad consequences Hájek lays out for it. More needs to be said here about how to formalize the idea that deliberation complicates prediction. But there is no longer and immediate and obvious worry that this position will mess up our formalisms. Nor is it clear that it puts new constraints on our credences, or entails counterintuitive consequences, like our rational-option credences poofing in and out of existence as we move in and out of deliberation (cf. Hájek’s section 4). If this is right, then maintaining that deliberation complicates prediction is far less costly than Hájek makes it out to be.

Suppose I am wrong, however, and there is no probability-gap free way of working out the idea that deliberation complicates prediction. This doesn’t spell doom for DARC, DCP, and their relatives. This is because, as I hope to have shown, there is a difficult and urgent problem about deliberation and prediction that such theses help us resolve. We may, then, consider accepting them, even if they mess up or force us to revise our formalisms.

5. Going forward.

Hájek launches a formidable attack on the idea that deliberation crowds out prediction. But if what I’ve said above is right, then there is a way of understanding this thesis so that it escapes some of the difficulties Hájek presents for it. Hájek may be right about some versions of this thesis, on which deliberation really does annihilate our rational credences. But those aren’t the most plausible and promising versions of the view.

Furthermore, we have a positive reason to accept some version of the claim that deliberation complicates prediction. Doing so explains the difference between the cases in which things are up to me and those in which they are not and helps resolve a difficult and urgent puzzle, which Hájek’s paper, for all it accomplishes doesn’t solve. Indeed, Hájek says nothing about this puzzle at all. This is important. If we are going to criticize
or reject a view, we should know what is at stake.

We are thus left with a new challenge: that of spelling out the details of the DCP thesis in a way that doesn’t lead to probability gaps—or, if that isn’t possible a way that resolves the puzzle so satisfactorily that it’s worth sacrificing our formalism. Much work thus remains, but it’s imperative that we do it. For the sake of our puzzle, as well as ourselves.¹⁵

¹⁵ For helpful comments and conversations, thanks to Maya Eddon, Alan Hájek, Sophie Horowitz, Berislav Marušić, Alejandro Pérez Carballo, and audience members at the Episteme conference in Phuket. Thanks especially to Jennifer Lackey, who organized the Episteme conference and to Alan Hájek for helpful conversations on his thought provoking paper.
References.


Hájek, A. “Deliberation Welcomes Prediction” in this volume.


