Deliberation Welcomes Prediction

Alan Hájek[1]

Tonight you will face the pleasant choice of having either red wine or white wine with your dinner. Perhaps you have subjective probabilities now for what you will end up choosing, or perhaps not. But a number of authors claim that once you begin *deliberating* about this choice, you *cannot* have such probabilities, on pain of irrationality. I disagree.

Let me approach this debate in stages.

1. Satellite view

Start with some propositional attitude. Are there propositions to which this attitude cannot be held, or at least cannot rationally be held? Are there *gaps* for the attitude, of necessity, or on pain of irrationality? I'll be especially interested in the (ir)rationality question; but of course, if there are propositions to which an attitude cannot be held *simplicter*, then it cannot rationally be held. It's especially interesting if there are *true* propositions to which the attitude cannot rationally be held.

For example, let the attitude be *knowledge.* Are there propositions—even true ones[2]—that cannot be known? According to Fitch's paradox, there are. Its upshot is that, necessarily, if in fact there is an unknown truth—and there surely is—then there is a truth that cannot be known. Or let the attitude be *belief.* Are there propositions that cannot rationally be believed? According to Moore's paradox, there are—propositions of the form

p & I believe that not *p*

or

p & I don't believe that *p*.

[1] For very helpful discussion and comments I thank especially Brad Armendt, Paul Bartha, J.C. Bjerring, Rachael Briggs, James Chase, John Collins, Nick DiBella, Kenny Easwaran, Edward Elliott, Caspar Hare, Brian Hedden, Yoaav Isaacs, Leon Leontyev, Isaac Levi, Hanti Lin, Aidan Lyon, Philip Pettit, Huw Price, Alex Pruss, Wolfgang Schwarz, Dan Singer, and Katia Vavova; and audiences at the University of Maryland Probability Workshop (organized by Aidan Lyon), the University of Pennsylvania, Oxford University, Arizona State University, Georgia State University, the Self-Prediction and Decision conference, Cambridge University (organized by Arif Ahmed and Huw Price), the *Episteme* conference, Phuket (organized by Jennifer Lackey), and the Australasian Association of Philosophy Conference, Macquarie University. Further thanks to Katia Vavova for her insightful comments on this paper (this volume).

2 Of course, false propositions cannot be known, since knowledge is factive.

Fitch's paradox and Moore's paradox give us, respectively, knowledge *blindspots* and belief *blindspots*, in Sorensen's (1988) term. He argues that there are more—e.g. in the surprise exam paradox ('I will receive a surprise exam this week') and propositions concerning the sharp boundaries of vague predicates.

Subjective probabilities—*credences*—are important, and perhaps even fundamental propositional attitudes. Are there rational credence blindspots—propositions to which one cannot rationally assign credences? As we will soon see, several authors think that there is an interesting and easily characterised class of such propositions.

2. Aerial view

I have long been interested in probability gaps. (My favourite examples have included non-measurable sets and chance gaps—see, for instance, my 2003.) If there are any such gaps for subjective probabilities, they are problematic for a number of pillars in Bayesian epistemology:

Conditional probability

I have campaigned for many years against the usual understanding of conditional probability as given by the ratio formula:

$P(A \mid B) = P(A \,\&\, B)/P(B)$, provided $P(B) > 0$.

(See especially my 2003.) One of my main arguments is from probability gaps. Let G be such a gap for P—that is, $P(G)$ is undefined. Nevertheless, various conditional probabilities involving G should be defined. For example, the conditional probability of G, *given* G, should be 1; and the conditional probability of not-G, *given* G, should be 0. The usual understanding founders on such cases, since the terms in the ratio are undefined.

Conditionalization

We may immediately parlay this argument against the ratio formula for conditional probability into an argument against *conditionalization*, which is standardly defined in terms of this formula. Conditionalization is the Bayesian's favored rule for updating credences. Suppose that initially your probability function is P. Now suppose that you learn a piece of evidence E, and that this is your total evidence. Your new probability function is supposedly given by:

For all X, $P_{new}(X) = P(X \mid E)$, provided $P(E) > 0$ (Conditionalization)

But the proviso fails if E was a probability gap for you before you learned it. We thus cannot model your becoming certain of such a gap by (Conditionalization). We will return to this thought later. (A similar point carries over to Jeffrey conditionalization, which models less decisive learning experiences in which one's probabilities across a partition are revised. Again, the ratio formula is usually assumed for the requisite conditional probabilities. See Jeffrey 1983.)

Decision theory

Decision theory's central notion is that of *expected* utility, a measure of the choiceworthiness of an option: the weighted average of the utilities associated with that option in each possible state of the world, the weights given by corresponding probabilities that those states are realized. The theory's mantra is that rational action maximises expected utility. Let G be a probability gap for you. Now consider these options:

Option 1: You get $1,000,000 if G, nothing otherwise.

Option 2: You get nothing.

You should clearly choose Option 1. But its expected utility is undefined, since the probability of G is undefined; so choosing Option 1 does *not* maximise expected utility.

Independence

Independence is a key concept in probability theory. When two propositions are independent, the truth of one of them is completely uninformative about the truth of the other, probabilistically speaking. Probability textbooks say that A and B are *independent* just in case

$P(A \,\&\, B) = P(A)P(B)$.

But what are we to say about probability gaps, for which the corresponding equation fails to hold? After all, if $P(G)$ is undefined, then $P(G)P(X)$ is undefined, for all X. If we say that propositions that are not independent are *dependent*, then we have the absurdity that G is dependent on every proposition. If we say that there is no verdict either way when any of the terms in the equation are undefined, then we will have silences where we want answers. For example, every proposition with probability less than one should turn out to be

dependent on itself (the truth of the proposition could not be more informative about its own truth!). But this account of independence cannot say this.

Regularity

Commonsense says, and many philosophers agree,[3] that some version of so-called *regularity* should hold:

If X is possible, then the probability of X is positive.

Versions of regularity as a constraint or condition on subjective probabilities have been proposed by Kemeny (1955), Shimony (1955, 1970), Jeffreys (1961), Edwards et al. (1963), Carnap (1963), Stalnaker (1970), Lewis (1980), Skyrms (1980), Appiah (1985), Jackson (1987), Jeffrey (1992), Benci et al. (2013). Of course, different senses of 'possible' give us different senses of regularity. But whichever sense we might favor, probability gaps will provide counterexamples. For if *G* is possible in the requisite sense, but the probability of *G* is undefined, then it is *not* the case that the probability of *X* is positive.

So it goes. Credence gaps wreak havoc for various foundational notions in probability theory, and thus in formal epistemology. So it is especially important if there is an interesting class of credence gaps for a rational agent.

And a number of prominent authors think exactly that. It is time to get in the trenches with them.

3. The 'deliberation crowds out prediction' thesis

Here is their view:

While deliberating about what you'll do, you cannot rationally have credences for what you'll do.

For example, suppose that you are deliberating about whether to have red or white wine with your meal. According to the view, you cannot rationally assign credences to these options. You cannot assign probability ½ to each of them; you cannot assign 2/3 to red and 1/3 to white (reflecting a partial inclination towards red); and so on. And so it is with all decisions, while you are in the process of deliberating about them. To be sure, *other* agents may be able to rationally assign credences to what you will choose. The problem is supposedly one of

[3] I disagree, for reasons given in my "Staying Regular?" (MS).

assigning credences *reflexively*—a 'first-person' problem. One cannot have *credences for one's own options while deliberating*—for short, *option-credences*—on pain of irrationality. In a slogan, the thesis is that *option-credences are irrational.*

This has come to be known as the 'deliberation crowds out prediction' thesis (following Levi 1997). At the risk of quibbling about terminology, I think that 'prediction' is normally used to connote a forecast or prophecy of a *specific* event. We often assign credences to things without making predictions (so understood). I assign credence 1/1,000,000 to each of the million tickets winning tomorrow's lottery, but I don't make any *predictions* regarding the winner. On the contrary—my uniform probability distribution reflects my *lack* of prediction. According to the thesis, deliberation doesn't crowd out just *prediction* (in the sense of a forecast, or a prophecy), but any assignment of credences to one's options. So I will give the thesis another name. I will call it *the DARC Thesis*: *D*eliberation *A*nnihilates *R*eflexive *C*redences.[4]

Versions of the DARC Thesis have been advocated by authors such as Spohn, Levi, Gilboa, Price, Louise, and others. Spohn and Gilboa couch their versions of the Thesis especially in terms of *modeling* credences: we should not model rational agents as assigning option-credences. Nevertheless, I take them to subscribe to the DARC Thesis. For example, Spohn (1977) writes: "probably anyone will find it absurd to assume that someone has subjective probabilities for things which are under his control and which he can actualize as he pleases" (115).

I don't.

4. Initial concerns about the DARC Thesis

Before examining various arguments for the DARC Thesis, let me register some concerns about it.

Firstly, when we are deliberating about what we will do, we are *uncertain* about what we will do; probability is our best theory of uncertainty; so offhand, we should assign probabilities to what we will do—or at least, we *may*. The Thesis is by no means a natural default position; quite the opposite.

[4] Thanks to Jessica Isserow for suggesting the 'A' in this acronym!

Secondly, the DARC Thesis imposes a new constraint on credences beyond the usual Bayesian ones that they obey the probability calculus and are revised by conditionalization. This is big news for Bayesianism, and if true, it should be written in neon lights. More than that, it seems to *contradict* orthodox Bayesianism. Consider a case in which initially you idly contemplate a decision that you will make some time in the future; time passes, and eventually the time to make the decision arrives, and you begin deliberating about it; then, finally, you have made your decision. It is perfectly compatible with the DARC Thesis that at the initial and final times, you have credences for what you will do; indeed, various proponents of the Thesis explicitly say that there is nothing problematic about such credences. It is only during the intermediate period while you are deliberating that the credences are forbidden. So suppose that you have credences for your options outside that period. Then we have *two* revisions of your probability function: your credences for your options suddenly *vanish* when your deliberation commences, and they suddenly *appear* when it ends. *Neither revision takes place by conditionalization*. This is inconsistent with the Bayesian orthodoxy that credences should be revised by conditionalization (or Jeffrey conditionalization). In fact, it is inconsistent with even the weaker claim that credences may be revised only on the basis of evidence (however such revision should take place). Not only are option-credences rationally permitted by these lights—they are rationally *required*.

Thirdly, and relatedly, the DARC Thesis conflicts with the Reflection Principle (van Fraassen 1984, 1995). According to Reflection, your credence in *X* at a time should be your expectation of your credences in *X* at any later time. Suppose that now you have not yet begun deliberating about the wine to accompany your dinner, and that you assign a credence to your choosing the red wine that evening—say, ½; no conflict with the DARC Thesis yet. Suppose also that you believe that you will be deliberating about the choice of wine at 7 pm. According to the DARC Thesis, your credence for red wine should be undefined then, so your expectation of your 7 pm credences for red wine should be undefined now. But your credence now is ½, not undefined, in violation of Reflection.

Fourthly, it seems clear to me that you can make *comparative* probability judgments about your options. 7 pm arrives, and you are in the thick of your wine deliberations. You judge that you are *more likely* to have red wine with your dinner than to have red wine with your dinner followed by port with your dessert, and this in turn is *more likely* than your having a shot of petrol. (The quirky restaurant you're at offers all these options.) And I don't see why numerical probabilities create any special problem. Comparative probabilities over your options may be especially apt when deliberation about them takes place over a long period of time, and things change during that time. Suppose that you are choosing between two job offers at two universities, but the decision is complicated, drawn out, and you must weigh many factors. Then the dean at one of the universities suddenly increases your salary—one of the factors—by $10,000. It is natural for your probability of taking that university's offer to suddenly be greater than it was, even if your deliberation is not yet over. This cannot be squared with the DARC Thesis. More generally, when you deliberate about your options, you weigh in turn the pros and cons of each of them. As a pro (con) regarding an option comes to your attention, it is natural for your probability of choosing it to increase (decrease).

Fifthly, consider a Perry/Lewis-style 'de se' ignorance case, in which you have a lot of evidence regarding an agent's decision situation, but you are uncertain who the agent is. You know a lot about their options, their utility function, their behavioral dispositions, and so on, and on that basis form credences about what they will choose. Then, you suddenly discover their identity: the agent is *you*! It seems odd that your credences should suddenly vanish.

Or suppose that there are two agents whose decision situations you are contemplating, and you know that one of the agents is *you*, but you are not sure who it is; you assign credence ½ to each of them being you.[5] (This is rather like the predicament of the Gods in Lewis's famous case (1979, 520-521).) According to the DARC Thesis, you cannot assign credences to both their choices, for in doing so, you will certainly assign a credence to your own choice. Yet by symmetry, you have no basis for assigning credence to one's choice but

[5] Thanks to Yoaav Isaacs for this case, and for those of the next paragraph.

not the other. That leaves only your assigning credences to *neither* of their choices. The DARC thesis ramifies, then, to your credences about the choices of agents besides yourself, somewhat surprisingly.

But this is only the beginning. Now suppose that you give *unequal* credence to each of the agents being you—say, 0.999 and 0.001. If you assign credences to the options of the agent who is *in fact* you, you sin against the DARC Thesis—even though you may have no way of figuring out this fact. This imposes an oddly externalist constraint on you. It seems that the only way that you can be sure of upholding the DARC Thesis is to withhold assigning credences to the options of *both* agents. The problem only gets worse when there are multiple agents who could be you, for all you know, however heavily your credences are biased towards one candidate. Indeed, such cases need not be far-fetched: when faced with a decision problem, and *any* uncertainty about whether that problem is yours or someone else's, this predicament is yours. But rationality may require of you at least *some* uncertainty about that! The DARC thesis ramifies still further.

Sixthly, and more generally, there will often be downstream consequences of the DARC Thesis for other propositions. I may believe or assign high credence to biconditionals between what I choose and other propositions. Suppose that I am deliberating between going to Italy and going to Iraq for a holiday, and I believe that my wife will be worried if and only if I go to Iraq. By the DARC Thesis, I cannot have a credence for 'my wife will be worried'—if I did, it would fix my credence for 'I go to Iraq' at the same value.

Or suppose that I believe a biconditional between something I am deliberating about now, and something I am not. I am choosing between having pizza and Chinese for dinner. Today is Tuesday, and Tuesdays are two-for-one pizza nights at my local pub. I believe that however I choose tonight, I will choose the same way next Tuesday. Indeed, I may believe this in virtue of my rationality: my reasons for one choice or the other will be the same then as now (we may suppose). By the DARC Thesis, I had better not have a credence for *next week's* choice.

Or suppose that I have a twin, faced with the same decision problem as I have, and I believe that however he chooses, so will I. By the DARC Thesis, I had better not have a credence for how *he* chooses. And these "sixthly" points carry

over even if there is merely a correlation between the relevant propositions, by my lights. For example, if I assign a credence to 'my wife will be worried', that will still constrain my credence in ‘I go to Iraq’; likewise for the later cases.

Similarly, probabilities propagate through entailments, so the DARC Thesis prohibits various credence assignments to consequences of one’s options. ‘I’ll go to Italy for a holiday’ entails ‘I’ll go somewhere for a holiday’. If I assign credence less than 1 to the latter, that puts an upper bound on my credence for the former. For example, if I assign probability 0.9 to my going somewhere for a holiday, then my probability for going to Italy for a holiday is bounded above by that value. In that case ‘I’ll go to Italy for a holiday’ is *not* a probability gap for me; at worst, my probability for it is an interval from 0 to 0.9. Then by the DARC Thesis, I must not assign credence less than 1 to ‘I’ll go somewhere for a holiday’, while deliberating about whether I’ll go to Italy.

To be sure, proponents of the DARC Thesis can embrace all these downstream consequences: indeed, they may say, I should not assign credence to my wife’s worrying, or to next Tuesday’s choice, or to my twin’s choice, or credence less than 1 to my going somewhere for a holiday. Still, we should be clear that the DARC Thesis ramifies to propositions about which I am not currently deliberating. Probability gaps are even more widespread.

Presumably it also ramifies to other propositional attitudes. Credence has various connections to other attitudes—belief, knowledge, and desire most obviously, but also hoping, wishing, dreading, fearing, and so on. I opened with a brief discussion of some examples of blindspots for knowledge and belief. Presumably the DARC thesis, if true, would furnish us with more blindspots, for these attitudes and others.

Be all that as it may, the DARC Thesis has much currency. If true, it would provide a rich source of probability gaps, in which I have indicated my long-standing interest. And there is nothing esoteric about these putative gaps—no fancy mathematics (such as the existence of non-measurable sets) or fancy metaphysics (such as the existence of chance gaps) is needed. Propositions such as ‘I will order red wine’ or ‘I will go to Italy for a holiday’ could hardly be more mundane. Yet if they are probability gaps, they are problematic for various foundations of probability theory and formal epistemology—recall my ‘aerial view’ of some of them in the previous section. The stakes are high.

5. Arguments for the DARC Thesis

I will now rehearse and rebut many of the main arguments for the DARC Thesis in the literature. I will group them according to their main themes: agency (arguments 1 – 4), vacuity (5), betting (6 and 7), and decision-theoretical considerations (8).

5.1 Wishy-washy agency/free will argument-sketches

I will begin with more of a gesture at an argument for the DARC Thesis than an argument *per se*. I have informally discussed the Thesis with many philosophers, a number of whom are sympathetic to it. When I press them for arguments, I often get a response that I will parody as follows:

Say some vague things about 'agency'.

Wave hands at 'free will'.

(Perhaps drop in a passing reference to Kant.)

Therefore, the DARC Thesis.

Even the published literature in favor of the Thesis has some moments like this. To be sure, there is something seductive about the idea that the very nature of free agency somehow precludes the assignment of rational credences to its upshots.

Needless to say, such a wishy-washy argument-sketch has to be filled in more before it is worthy of serious attention. But even in this inchoate form, we should be on the alert. For even this argument-sketch risks proving too much. The vague things said about 'agency' and the hands waved at 'free will' (and the passing reference to Kant) seem to apply just as readily in the third-person case as in the first-person case. When onlookers assign credences to my choices, they presumably regard me as an agent with free will. Yet there is not supposed to be anything problematic about such credences. Similarly, I may assign credences to my future choices when I am not deliberating about them—as it might be, next Tuesday's meal options—well aware that I will be a free agent then.

And I can also wave my hands, to the opposite conclusion. Free will may all the more ground one's option-credences. More on that shortly.

It is all to the good, then, that some authors fill in more this argument-sketch, giving arguments worthy of serious attention. I turn to them now.

5.2 There is no evidence for how an agent will choose from her own perspective (Price)

Huw Price (2007, 11) wants to distance himself from Pearl's view "that there is evidence [regarding how actors will choose] but that actors should ignore it". Price continues: "whereas the point should be that there isn't evidence in the first place, *from the actor's own perspective*." He is inspired by Ramsey (1978, 146): "any possible present volition of ours is (for us) irrelevant to any past event. To another (or to ourselves in the future) it can serve as a sign of the past, but to us now what we do affects only the probability of the future ... In a sense my present action is an ultimate and the only ultimate contingency."

Let me spell out what I take Price's argument to be in a bit more detail:

1. While deliberating, you have no evidence for how you will choose.
2. (Tacit) Your credences should reflect your evidence.

Conclusion: You should have no credence for how you will choose.

I have added premise 2, as some sort of bridge between evidence and rational credence (or lack thereof) is required.

I reply:

Premise 1

You typically *have* evidence regarding how you will choose—for example, your past choices in similar situations, or the choices of people similar to you in similar situations. Indeed, you typically have *better* evidence—not *worse*, let alone *none*—than onlookers do regarding your choice: the evidence of your own volition, too. Introspective evidence about your own volition provides *more* basis for the assignment of credences that you might use, beyond publically available information that others may unproblematically use. Following Joyce (2002), we may go so far as to say that your volition trumps (screens off) such information: it generates its own evidence. You get to make the news, and you get to know it ahead of everyone else. (More on this in §5.3.) It seems odd that *they* may assign credences to what you will choose, but *you* may not, when your evidential basis is superior to theirs.

We might even be able to say more about the nature of the extra evidence that you have at your disposal: for example, it might be a slight inclination to choose

one way rather than another, a 'tickle' that you can introspect, but that nobody else can detect.

Premise 2

Let's grant that your credences should reflect your evidence (in some sense); but they should *also* reflect your priors. And there is no difficulty with your assigning prior probabilities to what you will end up choosing, long before you face the moment of choice. In any case, perhaps option-credences are exceptions to premise 2; perhaps they give rise to a non-evidential way to have rational credences.[6]

The validity of the argument

Grant the premises for the sake of the argument. Does it follow that you should have no credence for how you will choose? Perhaps; but proponents of the Principle of Indifference will insist that, on the contrary, your lack of evidence *forces* certain credences on you.[7] To be sure, the Principle has hit hard times; for example, many authors think that Bertrand's paradoxes decisively refute it. Yet those same authors seem to happily reach for it when faced with the Monty Hall problem or the Sleeping Beauty problem—who would dare assign credence 1/17 to the prize being behind door #1 in the former problem, or allow Beauty to assign $1/\sqrt{2}$ to 'today is Monday' in the latter problem? We converge instead on answers dictated by the Principle of Indifference. The fact that the Principle fails to hold across the board does not show that it fails across the board. And certain cases of free choices may well be cases where it holds. In the middle of a maze, you must decide whether to go left or right. Perhaps when starting to deliberate about this, it is reasonable for you to assign probability ½ to each option.

Regarding the quote from Ramsey: I don't see there an argument for the DARC Thesis. I take the closest point in the neighborhood to be that evidence from the past that may be available to others is unavailable to me:[8] for example,

---

6 Thanks to Yoaav Isaacs for the first point, and to Wolfgang Schwarz for the second.

7 If we regard the Principle of Indifference as a constraint on your priors, then we may subsume this point under my discussion of Premise 2.

8 Even this reading is not immediate. I take it that if any possible present volition of ours is (for us) irrelevant to any past event, then by the symmetry of irrelevance—independence—any past event is irrelevant to any possible present volition of ours (for

from my past behaviour, or from the past behaviour of similar people in similar situations. But that falls short of showing that I have *no* evidence; indeed, following Joyce, I may have even *better* evidence, of a different kind. In any case, I think that Ramsey would have been happy assigning subjective probabilities even in the absence of evidence. That's exactly what priors are!

5.3 The agent's authority trumps other evidence (Price)

Price (2012, 529) writes: "One route to [the DARC] thesis … turns on the special epistemic authority of the deliberating agent, concerning her own actions. This authority 'trumps' any merely predictive knowledge claim about the same matters, rendering it necessarily unjustified."

I reply: Grant that the deliberating agent has special epistemic authority regarding her own actions. Grant that this authority even 'trumps' any evidence that onlookers might use to ground a credence regarding what she will do. (That is how I understand the second sentence. "Knowledge claim" seems a strange turn of phrase here—we may assume that nobody *knows* what the agent will do, neither her nor others, in the cases of most interest to us.) But if anything, this means that it's the *onlookers* who should lack the credences regarding what she will do, not the agent herself! After all, they lack the crucial information that the agent has—the trump card—which grounds her special epistemic authority. And the agent may have such information well before she has decided what she is going to do—for example, the initial tickle that she is inclining slightly towards the red wine is evidence that only she has.

The DARC Thesis seems to be *counter*-supported rather than supported by this argument. Having special epistemic authority about a proposition is *more* reason to assign it a credence, not less.

5.4 Incompatible presuppositions about agency (Louise)

Jennie Louise (2009, 339) canvases the following argument for the DARC Thesis. While she does not explicitly endorse it, she does seem to regard it as "a

---

us). Then information about the past does not provide evidence (for us) for what we will choose. But if Ramsey had intended to say that, it is surprising that he put the point about irrelevance the other way round.

real problem … if we think of agency as an all-or-nothing affair" (340). Her response is to "think of agency as something that comes in degrees"; by contrast, I do not regard it as a real problem even if agency is all-or-nothing.

> ... it might be argued that while there is no logical incompatibility between prediction and deliberation, there is nevertheless a pragmatic contradiction—a Moorean paradox… Deliberation about whether to U … presupposes agency: that is, it presupposes that I am the kind of creature who can be motivated to act in accordance with my perceived reasons. Prediction that I will U … on the other hand, seems to presuppose the opposite: that I am merely ‘‘an object in the world’’ subject to causal forces outside my control. Using self-prediction in deliberation therefore involves the pragmatic contradiction of assuming both that one is an agent and that one is not an agent.

I reply: Assuming both that *one is* an agent and that *one is not* an agent involves a *semantic* contradiction rather than a pragmatic one, so calling this a “Moorean paradox” is an understatement. Rather, the point seems to be that any agent who has option-credences must make inconsistent presuppositions. I offer the following reading of this argument:

1. Deliberation about what one will do presupposes one's agency.
2. Predicting and assigning credences to what one will do presupposes one's non-agency.

Hence,

3. Option-credences require one to make incompatible presuppositions, which is irrational.

Conclusion: The DARC Thesis.

I baulk at 2. Again, it unhappily generalizes to onlookers: *their* assigning credences to what I will do presupposes my non-agency. But why? Why must they regard me as merely “an object in the world” subject to causal forces outside my control? On the contrary, they may well have a far better understanding of the causal forces *under* my control *in virtue of my agency*—for example, by mentally putting themselves in my shoes. And 2 unhappily generalises to my credences about what my future self will do even when I am *not* deliberating about it—unhappily for similar reasons. Once we see all this, we also see how it carries over to my own case when I am in the midst of deliberation. There is no contradiction—semantic, or otherwise—in my

assigning credences to what I will do *in virtue of my agency*.

* * * * * * * * * * *

For all that I have said here and in the previous few sections, 'agency'-based arguments have a grip on various of my interlocutors. A common thought is that deliberation requires one to take a first-person perspective on oneself, while option-credences require one to take a third-person perspective on oneself. One cannot adopt both perspectives at the same time, although one may adopt one perspective or the other at a given time. It's rather like the way one cannot see the duck-rabbit both 'as' a duck and 'as' a rabbit at the same time, the thought goes, although it can be seen one way or the other at a given time. Perhaps this thought can be dispelled somewhat by noticing that talk of 'perspectives' or 'seeing as' may be misleading here, suggesting as it does one particular occurrent attitude, or phenomenological state, and then another. But like beliefs, credences need not be occurrent. You can *have* a credence without attending to it or being aware of it.[9]

And there may not be any associated phenomenological state at all. We often regard *groups* as agents, and we often attribute beliefs and even credences to them. Groups also make decisions, and deliberate about them. Presumably the DARC Thesis should hold for group agents too, if it holds at all. But such agents do not have phenomenological states. Indeed, the Thesis seems especially implausible for such agents[10], perhaps partly for that reason.

I think that considerations of agency do not support the DARC Thesis, at least as far as we have seen so far. But they may support something in the neighborhood. More recently, Price has suggested that while deliberating, an agent may have credences for what she will choose, but such credences are *muted*—that is, they do no play their usual functional role. In particular, they do not play their usual role in guiding her decision.[11] They are disengaged from her action; deliberating is like a mental 'clutch' with respect to her option-credences. That said, it may be difficult to square this view with a functionalist account, according to which a credence is *identified* by its functional role. And even

9 Thanks here to Philip Pettit for helpful discussion.

10 I am grateful to Aidan Lyon for this point.

11 Price presented this idea in his talk "What Ramsey Got Right", at the *Self-Prediction and Decision* conference, Cambridge University, May 2015. I thank him for very helpful discussion on this point, and John Cusbert for the "clutch" metaphor that follows.

without subscribing to functionalism, I would like to hear more about what option-credences *are* when they are muted. Furthermore, later we will see quite different pictures of deliberation, due to Skyrms and Arntzenius, which portray option-credences as central to decision-making; indeed, deliberation is largely *constituted by* the updating of option-credences! We will also see how game theory—a theory of decision-making par excellence—appeals to them. So for various reasons, we may need them to speak, and speak loudly.

All that said, muted credences are still credences. So while this proposal may be in the neighborhood of the DARC Thesis, it is a retreat from the DARC Thesis itself—which is all to the good.

5.5 The argument from vacuity (Levi)

Isaac Levi is a pragmatist—fittingly for the John Dewey Professor of Philosophy Emeritus at Columbia University. The centerpiece of his pragmatism is this idea: "Changes in point of view are to be justified by showing that they are optimal or, more generally admissible, as options among the available options and relative to the goals of the inquiry" (2007, 4). Nobody has defended the DARC Thesis in more detail or more famously. It behoves me to present and respond to a chief argument of his in detail.

To understand Levi's argument, we need a bit of background. Suppose that an agent has to choose from a set of options A, and she fully believes that she will implement exactly one of them. An 'admissible' option is one that is not ruled out by the principles of choice. (Maximizing expected utility is one such principle.) 'Serious possibilities' are propositions that are compatible with her full beliefs. The 'weak thesis' is that she should judge every hypothesis that says that a specific option in A is implemented to be a serious possibility.

Now, the argument. Here is a key passage in Levi's (2007, 4-5).

> … even on the assumption that credal probabilities are used to compute expected values of options and, hence, to determine fair betting rates, the argument I offer does not show that the decision maker's assigning unconditional probabilities to hypotheses concerning what the decision maker will do is incoherent *unless one wants the standards of rationally coherent probability judgments to serve in the evaluation of expected value in contexts of choice where the principles of choice will sometimes be applicable non vacuously*. I do want to focus on probability judgment on the assumption that principles of choice can be applied non vacuously to the determination of which of the available options are admissible and which are not. But my

argument does not show that rejecting this assumption is incoherent or inconsistent.

I show that a rational decision maker assigning probabilities to hypotheses concerning the option he is about to choose must assign full belief to the prediction that the decision maker will choose rationally (i.e., choose an admissible option among those judged to be available) on pain of incoherence or contradiction. Only admissible options are seriously possible according to the decision maker. From this, according to the weak thesis, the admissible options and the feasible or available options must coincide. Hence, criteria for rational choice must be vacuously applicable in *all* contexts of choice.

If this vacuous applicability thesis is rejected, the admissible and the available cannot coincide in all contexts of choice. Hence, the criteria for rational probability judgment should prohibit a rational decision maker from assigning unconditional credal probabilities to hypotheses concerning which available option will be chosen *on pain of inconsistency with the vacuous applicability thesis*.

Let me present what I take the argument to be in 'premise, premise, … conclusion' form:

1. The criteria for rational choice apply non-vacuously in some choice contexts.
2. But when the admissible options and the available options for an agent coincide, the criteria for rational choice apply vacuously.

Hence,

3. There are choice contexts for an agent in which the admissible options and the available options for her do *not* coincide.
4. Suppose that there is a rational agent in such a choice context, and suppose for reductio that she assigns credences to her choices.
5. She must be certain that she will choose rationally.

   Hence,

6. Only admissible options are seriously possible for her.

   Hence,

7. The admissible options and the available options coincide for her.
8. This contradicts our supposition in 4.

   Hence,

   Conclusion: no rational agent assigns credences to her choices while deliberating about them.

I reply:

Premise 2

Consider a case in which all the agent's options are equally good, and so are tied by the lights of her criteria for choice—for example, she has two options, tied in expected utility. (The choice between red and white wine could easily be like that; it often is for me.) The admissible options and the available options for the agent coincide. But the criteria for rational choice apply *non-vacuously*: she still faces a genuine choice. She may, for instance, use a fair coin to make her choice, and as a result she assigns credence ½ to each of them (contra the DARC Thesis); indeed, she is rationally *required* to do so if she knows that the coin is fair.

Premise 5

Rational agents who assign credences to their options need not be *smug*: they need not be certain that they are rational. On the contrary, rationality may even *require* them to give positive credence to their *not* being rational. After all, their being rational (or not) is in each case a contingent matter, and as such their credence in it should be appropriately responsive to their evidence. Among other things, they should realise that they may have misleading evidence regarding their rationality. In any case, their evidence is typically not so conclusive as to eliminate even the tiniest crack of doubt about their rationality. (Presumably, that evidence gives at best inductive support to their rationality.) Moreover, even supposing that a rational agent is certain that she is rational *now*, she need not be certain that she will *stay* rational—typically, her evidence will not settle that either. In particular, she might assign some positive credence to her not choosing rationally when the moment of choice comes. Indeed, she might not even be certain that she will choose at all!

The inference from 5 to 6

Grant for the sake of argument that the agent is certain that she will choose rationally (when the time comes). But she may nonetheless deliberate: *to figure out what the admissible options are*! She is certain that whatever she eventually chooses will be admissible, but it takes her a while to determine what that will be. And during that time, some *inadmissible* options are also serious possibilities for her; they remain so until she rules them out. So I question the inference here

from her being certain *that* she will choose rationally to her being certain *what* she will rationally choose. And even after she has figured out what the admissible options are, she may still want to deliberate—to *check* what she has figured out.

Perhaps Levi could reply: "I was assuming that the agent is ideally rational, so she takes no time to figure out what the admissible options are—she sees them immediately". But then she does not deliberate! Then the thesis isn't really 'deliberation crowds out prediction (for rational agents)'; instead, it's 'rationality crowds out deliberation'![12]

The inference from 6 to 7

The natural reading of an option's being "available" for her is that she is *able* to perform it. Levi uses the word "feasible" interchangeably with "available", which only reinforces this reading to my ear. But then 7 does not follow from 6. Even when you are certain that you *will not* perform a given action, you may well be *able* to perform it. Even when you are certain that you *will* choose the red wine, you are still *able* to choose the white. (And even a shot of petrol, for that matter.)

Levi must be using the word "available" in a doxastic sense: an option is "available" if your choosing it is compatible with your full beliefs. In that sense, the white wine is no longer "available" once you have decided to choose the red (and perhaps the shot of petrol never was). But then premise 2 seems less plausible to me, for another reason. When you have made up your mind not to choose various options, so that they're "unavailable" in Levi's sense—even though you are still *able* to choose them—the criteria for rational choice may still apply *non*-vacuously, confirming that you're doing the right thing. (For example, drinking the red wine really has higher expected value than the white for you after all.)

The validity of the argument

[12] Cf. Stalnaker (1991, 428-429): "Deliberation, to the extent that it is thought of as a rational process of figuring out what one should do given one's priorities and expectations is an activity that is unnecessary for the deductively omniscient. In fact any kind of information processing or computation is unintelligible as an activity of a deductively omniscient agent."

This argument (and indeed various other arguments for the DARC Thesis) does not rule out one's having credences for *some but not all* of one's options. This seems especially to be a live possibility when one has several options. Why can't one have a *partial* credence function, defined on some of the options but not others? Yet the conclusion of Levi's argument rules that out too, and it is thus stronger than the argument for it supports, as far as I can see.

In any case, the argument does not support the full DARC Thesis ("no rational agent assigns credences to her choices while deliberating about them")—just the restriction of it to choice contexts for an agent in which the admissible options and the available options for her do *not* coincide. As we have seen, there are contexts in which they *do* coincide—even non-vacuous ones, in which all one's options are equally good.

Finally, I wonder how to square the DARC Thesis with Levi's pragmatism—recall its centerpiece in the "Changes in point of view" quote, above. Consider a case in which you have credences for a choice further in the future, but you are not yet deliberating about it—say, next Tuesday's meal. When you begin deliberating about that choice, your credences are supposed to vanish according to Levi. And having completed the deliberation, the credences may reappear (presumably as 1 to what you will actually do and 0 to everything else). That involves *two* "changes in point of view". But I wonder how you—or Levi—could justify them by showing that they are "optimal or, more generally admissible".

5.6 Betting rates collapse (Levi)

In another of his arguments for the DARC Thesis, Levi (2007) identifies credences with fair betting rates. For example, your credence that it will rain tomorrow is 0.7 if and only if 70 cents is your fair betting price for a bet that pays $1 just in case it rains.

Levi turns to the specific problem of an agent having a choice between two options, **a** and **b**, with **a** strictly preferred to **b**; he imagines that she has credences for these options, *and* that she uses these credences to determine fair betting rates for bets on whether **a** or **b** will be chosen. He shows that the only admissible option is to choose **a** and to regard the fair betting rate for her doing

so to be 1 (and 0 for choosing **b**). Fair betting prices for one's options collapse to 0 and 1.

I reply: The problem is not with the agent assigning credences to her choices, but rather with the identification of credences with fair betting rates. This identification founders especially on cases in which the placement of the bet interferes with the proposition bet on. For example, I am highly confident that I will fall asleep tonight: I do so every night (eventually). But now suppose that I have placed a bet that will pay a million dollars just in case I fall asleep tonight. That very fact will make my credence that I will fall asleep tonight drop sharply. I picture myself at 5 a.m. pounding my pillow, working myself up into a panic: "Quick! Fall asleep! *Fall asleep!!*" And betting on which option one will choose is another such case—it interferes with the proposition bet on.

In fact, I take it that elsewhere Levi himself does not subscribe to the identification of credences with fair betting rates. He rejects the diachronic Dutch Book argument for conditionalisation, according to which a non-conditionaliser will be susceptible to sure losses on bets that she considers fair. Levi (1988) replies that the agent will see the sure loss coming, and so will not take all the requisite bets. But that is to reject the betting analysis.

5.7 Betting rates cannot be applied (Spohn)

I mentioned earlier that Spohn's version of the DARC Thesis focuses on the credences for acts in *decision models*, and soon I will countenance an argument of his that is explicitly about such models. However, as I also mentioned, I take him also to be a proponent of the DARC Thesis itself—an early and important one—and so do other authors (e.g., Rabinowicz 2002). Moreover, some of Spohn's arguments are not confined to considerations about decision models. For example, in his (1977, 115) he writes:

> It is generally acknowledged that subjective probabilities manifest themselves in the readiness to accept bets with appropriate betting odds and small stakes. Hence, a probability for an act should manifest itself in the readiness to accept a bet on that act, if the betting odds are high enough. Of course, this is not the case. The agent's readiness to accept a bet on an act does not depend on the betting odds, but only on his gain. If the gain is high enough to put this act on the top of his preference order of acts, he will accept it, and if not, not.

I understand the argument as follows:

1. The attractiveness of a bet on one's act depends only on the difference between the bet's prize and its price. [This can be shown with a bit of work—see Rabinowicz (2002), which has helped my understanding of this argument.]

In particular,

2. The bet's attractiveness does not depend on the agent's fair betting rate.

Hence,

3. The notion of a fair betting rate cannot be applied to a bet on one's act.

4. Credences are fair betting rates.

Hence,

Conclusion: The DARC Thesis.

I reply:

Nowhere does *deliberation* enter into this argument. So it should apply equally to your future acts that you are not deliberating about. Be that as it may, again we should reject the identification of credences with fair betting rates, premise 4.

5.8 There is no decision-theoretic role for option-credences (Spohn)

Finally, I turn to an argument of Spohn's (1977) that explicitly concerns option-credences in decision models, but that implicitly supports the DARC Thesis itself; indeed, I regard it as one of the most interesting arguments for the Thesis. (Levi 2007 has a similar argument, with acknowledgment to Spohn.)

Spohn writes:

> probabilities for acts play no role in decision making. For, what only matters in a decision situation is how much the decision maker likes the various acts available to him, and relevant to this, in turn, is what he believes to result from the various acts and how much he likes these results. At no place does there enter any subjective probability for an act. The decision maker chooses the act he likes most – be its probability as it may. But if this is so, there is no sense in imputing probabilities for acts to the decision maker. For one could tell neither from his actual choices nor from his preferences what they are. Now, decision models are designed to capture just the decision maker's cognitive and motivational dispositions expressed by subjective probability and utilities which manifest themselves in and can be guessed from his choices and preferences. Probabilities for acts, if they exist at all, are not of this sort, as just seen, and therefore should not be contained in decision models. (115).

The key premise here is:

Probabilities for acts play no role in decision making.

The explicit conclusion here is:

Conclusion: Probability for acts should not be contained in decision models.

The final sentence seems to leave open the possibility that probabilities for acts exist, as far as this argument goes. However, I take it that Spohn intends to close off that possibility with his betting argument (see §5.7 above), which concerns credences themselves—rather than decision models—and which immediately follows this argument in his paper.[13] And recall the quote in which he expresses such scepticism for the notion of credences for things under one's control as to call it "absurd".

Regarding the key premise, I reply: Probabilities for acts play no role in *Savage's* decision theory, but they do in some other decision theories. Taken at face value, Jeffrey's (1983) evidential decision theory requires probabilities for acts. After all, its formula for expected utility of an act A is a weighted average of utilities associated with A across a partition of states $\{S_i\}$, the weights given by conditional probabilities of the form $P(S_i \mid A)$. And he understands these conditional probabilities in terms of the usual ratio formula: $P(S_i \,\&\, A)/P(A)$. To be sure, I would prefer understanding them as primitive conditional probabilities instead. Be that as it may, the usual reading of evidential decision theory has a role for option-credences.

It seems that causal decision theory also requires option-credences. Consider a case in which you take propositions about what you will do to be correlated with dependency hypotheses (Lewis 1981), or their analogues in other versions of causal decision theory. (See Gibbard and Harper 1978 and Egan 2007 for such cases.) Then, if you don't have option-credences, you cannot have credences in the relevant dependency hypotheses, and hence cannot assign causal expected utilities to the relevant options.[14]

Moreover, probabilities for acts play a crucial role in Arntzenius's (2008) 'deliberational decision theory', which is inspired by Skyrms's 'deliberational dynamics' (1990)—more on those shortly. And Hare and Hedden (forthcoming)

---

[13] I have reversed their order so that Spohn's betting argument immediately followed Levi's above, and to provide a natural segue to the theoretical roles for option-credences below.

[14] Thanks here to Brian Hedden.

argue that in certain decision problems, one needs to assign low credence to one's taking a manifestly bad option in order to get the right results. To be fair, Spohn's paper predates these works. Still, they may be regarded as existence proofs that option-credences may play a central role in decision models after all.

In any case, credences have other theoretical roles besides their role in decision models. Indeed, I would have thought that their primary role is to *codify uncertainty*. That uncertainty may be parlayed into rational decision-making, but it is more fundamental. It may be parlayed elsewhere, too—for example, into confirmation relations, inference, and in connections to binary belief (for example, via the Lockean Thesis, or some other bridge between credence and belief). In short, even if probabilities for acts play no role in *decision theory*, that doesn't mean that they play no useful theoretical role. Soon we will see more such roles that they play.

Furthermore, even if option-credences play no role in decision theory, it does not follow that they do not *exist*. Edward Elliott (MS, 6) puts the point well, arguing that this form of reasoning proves too much (his discussion concerns Savage's decision theory):

> The strongest argument for [the DARC] thesis seems to be that credences regarding which action will be chosen in the present circumstances play no role in rational decision-making … Even supposing that this is true, it is quite different than the thought that we simply do not have such credences. Utilities for events also play no role in decision-making either, according to the orthodoxy—by hypothesis, what event obtains is independent of the choice between acts, so any valuation of the events on the subject's behalf is regarded as irrelevant to her choice. It would be unreasonable to infer from this that we do not have utilities for events; at least, it certainly seems to me that I am able to judge which of two events I would rather be true, even if I know that this is entirely beyond my control.

Coming from the other side, there are also theoretical *costs* to the DARC Thesis's prohibition of option-credences. It introduces a new necessity[15]—the avoidance of option credences, on pain of irrationality—and this complicates our epistemology and decision theory. Of course, some theoretical costs are worth paying—one should not cut oneself on Ockham's razor. For all I have said, the distinction may earn its keep. Still, the case for the DARC Thesis from theoretical considerations must take all this into account.

There is also the problem that the DARC Thesis is in tension with other putative rationality principles. I have already mentioned its tension with conditionalization (and Jeffrey conditionalization), with the Reflection Principle (§4), and with the Principle of Indifference (§5.2). Moreover, according to the Principal Principle, your credence in a proposition, conditional on the chance of that proposition having a particular value, should be that value.[16] Imagine that you learn the *chance* that you will choose a particular option while deliberating about it—this seems like a coherent possibility. By the Principal Principle, you are *required* to make that your credence, and in particular required *not* to have a credence gap for that option. The DARC Thesis and the Principal Principle then collide—and you can guess which one *I* side with!

More generally, following Gaifman (1988) we call Q an *expert function* for you if your credence function P is constrained by:

$$P(X \mid Q(X) = p) = p \text{ (provided } P(Q(X) = p) > 0).$$

Roughly, you align your credences with what you take your expert's probabilities to be. Chance is just one expert function for you. According to the

---

[15] I thank James Chase for this way of putting the point.

[16] That's not quite what the Principle says (see Lewis 1980 for its proper statement), but close enough for my purposes here.

Reflection Principle, your credence function at a future time is another. According to Rational Reflection (Christensen 2010), the evidential probability function is another (when conditionalised on your evidence). Or you may have more specialised experts—for example, a metereologist with respect to the probability of rain. The DARC Thesis collides with your knowing the probability assignment of any expert for you regarding what you will do, while you are deliberating about that. For by your own lights, you are supposed to align your probability assignment with that of your expert, while the DARC Thesis forbids you to have any assignment at all.

But this still understates the case for allowing option-credences. It isn't just that forbidding them may have some unwanted theoretical consequences; allowing them has some genuine theoretical benefits. I believe that they earn their theoretical keep, and then some.

6. Theoretical roles for option-credences

I have already mentioned some theoretical benefits of positing option-credences. That's just the start.

Option-credences also play an important role in game theory. They are appealed to in the notion of *correlated equilibrium*, at least as it is standardly presented. A more general solution concept than that of Nash equilibrium, we are to imagine each player choosing her action on the basis of a signal that is observed by all the players. Common knowledge is assumed of all the players' probabilities for performing each of their possible actions; this includes each player's probabilities of performing her *own* possible actions.

Trembling hand equilibrium is a refinement of Nash equilibrium that takes into account the possibility of off-equilibrium play, through an unintentional slip or 'trembling hand'. The players' probabilities of such actions are supposed to be negligible, but again common knowledge; and again, they include each players' probabilities of her *own* trembles.

Option-credences are also central to Skyrms's (1990) *deliberational dynamics.* The idea is that deliberation is a dynamic process in which an agent repeatedly revises her credences in her options. She begins in a state of indecision, with a set of credences over her options. She calculates their expected utilities on the basis of those credences. She then increases her credence in each option whose

expected utility is greater than the status quo. And so it goes, her option-credences evolving until at the moment of decision the probability for the chosen action becomes virtually 1. Skyrms uses deliberational dynamics to model phenomena as diverse as the arms race, the emergence of conventions, and the social contract. Bartha (2012) uses it to model one's evolving credences when faced with many theistic hypotheses, à la the 'many Gods' objection to Pascal's Wager. And Bartha, Barker, and Hájek (2013) apply it to 'Satan's apple', a decision puzzle due to Arntzenius, Elga and Hawthorne (2004).

Arntzenius's (2008) *deliberational decision theory* is inspired by Skyrms's deliberational dynamics. He understands a 'mixed decision'—a decision to perform one's possible actions with specified probabilities—as having those probabilities at the end of rational deliberation. The agent iteratively updates her option-credences, Skyrms-style, until eventually they reach an equilibrium—though they may not be anywhere near extremal at that point. So, far from deliberation crowding out option credences, it is largely *constituted by* such credences. Indeed, deliberation is characterised in terms of them: it terminates when they reach an equilibrium. Deliberational decision theory can be given both an evidential and a causal formulation. If the expected utilities are causal, we have 'deliberational causal decision theory'; if they are evidential, we have 'deliberational evidential decision theory'. Arntzenius advocates deliberational causal decision theory, and he argues that it can solve some apparent 'counterexamples' to causal decision theory.

Both Skyrms and Arntzenius appeal to deliberational dynamics to unify decision theory and game theory. Skyrms notes the somewhat puzzling fact that while decision theorists insist that rational decision-making is always a matter of maximizing expected utility, game theorists cast it instead in terms of various equilibrium concepts (Nash, correlated, trembling hand, and so on). He shows that by taking deliberation to be a dynamic process, expected utility maximization can be taken as fundamental, and the various game-theoretic solution concepts as derived from it—they correspond to different kinds of dynamic stability. Now *that's* a powerful theoretical role for option-credences!

7. Conclusion

I like gaps. I am intrigued by the idea that there can be blindspots for various propositional attitudes. I have championed the existence of gaps for rational credences, and I have marshaled them to argue that various basic probabilistic notions need to be rethought. This, in turn, forces us to rethink important parts of formal epistemology.

The DARC Thesis, if true, would provide me with more ammunition for my dark theses about the current state of play in these areas. Probability gaps would be as mundane as choosing red wine over white, or a holiday to Italy over one to Iraq. They would arise every time a rational agent deliberates about what to do. But I can't in good conscience help myself to these putative credence gaps. Not only am I unconvinced by the arguments for the DARC Thesis; I believe there are strong reasons to reject it outright.

Now that I'm done with this paper, I'll reward myself with a glass of wine. Now that you've finished reading it, you might want to do so too. I'm highly confident, but not certain, that I'll choose the red. How about you?

School of Philosophy
Australian National University

REFERENCES

Appiah, Anthony (1985): *Assertion and Conditionals*, Cambridge: Cambridge University Press.

Arntzenius, Frank (2008): "No Regrets, Or: Edith Piaf Revamps Decision Theory", *Erkenntnis* 68 (2), 277-297.

Arntzenius, Frank, Adam Elga and John Hawthorne (2004): "Bayesianism, Infinite Decisions, and Binding", *Mind* 113, 251-283.

Bartha, Paul (2012): "Many Gods, Many Wagers: Pascal's Wager Meets the Replicator Dynamics", in *Probability in the Philosophy of Religion*, eds. Jake Chandler and Victoria Harrison, Oxford: Oxford University Press.

Bartha, Paul, John Barker, and Alan Hájek (2013): "Saint Peter, Saint Petersburg, and Satan" *Synthese* 191 (December), 629-660, DOI 10.1007/s11229-013-0379-9.

Benci, V., L. Horsten, and S. Wenmackers (2013): "Non-Archimedean Probability (NAP)", *Milan Journal of Mathematics*, 81, 121-151.

Carnap, Rudolf (1963): "Carnap's Intellectual Autobiography", in *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, Vol. XI, ed. Paul Arthur Schilpp.

Christensen, David (2010): "Rational Reflection", *Philosophical Perspectives* 24 (1), 121-140.

Edwards, W., Lindman, H., and Savage, L. J. (1963): "Bayesian Statistical Inference for Psychological Research", *Psychological Review* 70, 193-242.

Egan, Andy (2007): "Some Counterexamples to Causal Decision Theory", *Philosophical Review* 116 (1): 93-114.

Gaifman, Haim (1988): "A Theory of Higher Order Probabilities", in *Causation, Chance, and Credence*, Brian Skyrms and William L. Harper (eds.), Dordrecht: Kluwer Academic Publishers.

Gibbard, Allan and William Harper (1978): "Counterfactuals and Two Kinds of Expected Utility", in C. A. Hooker, J. J. Leach, and E. F. McLennen (eds.), *Foundations and Applications of Decision Theory*, Vol. 1, Dordrecht: D. Reidel.

Gilboa, Itzhak (1994): "Can Free Choice Be Known?", in Christina Bicchieri, Richard Jeffrey and Brian Skyrms (eds.), *The Logic of Strategy*, Oxford: Oxford University Press.

Hájek, Alan (2003): "What Conditional Probability Could Not Be", *Synthese* 137, No. 3, December 2003, 273-323.

Hájek, Alan (MS): "Staying Regular?".

Hare, Caspar and Brian Hedden (forthcoming): "Self-Reinforcing and Self-Frustrating Decisions", *Noûs*.

Jackson, Frank (1987): *Conditionals*, Oxford: Blackwell.

Jeffrey, Richard C. (1983): *The Logic of Decision*, 2nd edition, Chicago: University of Chicago Press (first edition: 1966).

Jeffrey, Richard (1992): *Probability and the Art of Judgment*, Cambridge: Cambridge University Press.

Jeffreys, Harold (1961): *Theory of Probability*, 3rd ed., Oxford: Clarendon Press.

Joyce, James M. (2002): "Levi on Causal Decision Theory and the Possibility of Predicting One's Own Actions", *Philosophical Studies* 110 (1): 69 – 102.

Kemeny, J. (1955): "Fair Bets and Inductive Probabilities", *The Journal of Symbolic Logic* 20 (3), 263–273.

Levi, Isaac (1988): "The Demons of Decision", *The Monist* 70: 193-211.

Levi, Isaac (1993): "Rationality, Prediction, and Autonomous Choice", *Canadian Journal of Philosophy* 23 (sup.1), 339-363.

Levi, Isaac (1997): *The Covenant of Reason: Rationality and the Commitments of Thought*, Cambridge: Cambridge University Press.

Levi, Isaac (2007): "Deliberation Does Crowd Out Prediction", in *Hommage à Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz,* eds. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonssson. www.fil.lu.se/hommageawlodek

Lewis, David (1979): "Attitudes De Dicto and De Se", *The Philosophical Review* 88, No. 4, 513-543.

Lewis, David (1980): "A Subjectivist's Guide to Objective Chance", in *Studies in Inductive Logic and Probability*, Vol II., ed. Richard C. Jeffrey, University of California Press.

Lewis, David (1981): "Causal Decision Theory", *Australasian Journal of Philosophy* 59 (1), 5 – 30.

Louise, Jennie (2009): "I Won't Do It! Self-Prediction, Moral Obligation and Moral Deliberation", *Philosophical Studies* 146 (3):327 – 348.

Price, Huw (2007): "The Effective Indexical", http://philsci-archive.pitt.edu/4487/.

Price, Huw (2012): "Causation, Chance, and the Rational Significance of Supernatural Evidence", *Philosophical Review* 121 (4), 483-538.

Rabinowicz, Wlodek (2002): "Does Practical Deliberation Crowd Out Self-Prediction?", *Erkenntnis* 57 (1), 91-122.

Ramsey, F.P. (1978): "General propositions and causality", in D.H. Mellor (ed.), *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, London: Routledge and Kegan Paul,133-151.

Shimony, Abner (1955): "Coherence and the Axioms of Confirmation", *Journal of Symbolic Logic* 20, 1-28.

Shimony, Abner (1970): "Scientific Inference", in Robert Colodny (ed.), *The Nature and Function of Scientific Theories*, University of Pittsburgh Press.

Skyrms, Brian (1990): *The Dynamics of Rational Deliberation*, Cambridge, MA: Harvard University Press.

Skyrms, Brian (1995): "Strict Coherence, Sigma Coherence, and the Metaphysics of Quantity", *Philosophical Studies* 77, No. 1, 39-55.

Sorensen, Roy (1988): *Blindspots*, Oxford: Oxford University Press.

Spohn, Wolfgang (1977): "Where Luce and Krantz Do Really Generalize Savage's Decision Model", *Erkenntnis* **11**, 113–134.

Stalnaker, Robert C. (1970): "Probability and Conditionals", *Philosophy of Science* 37, 64-80.

Stalnaker, Robert (1991): "The Problem of Logical Omniscience, I", *Synthese* 89 (3), 425 - 440.

van Fraassen, Bas (1984): "Belief and the Will", *Journal of Philosophy* 81 (5):235 - 256.

van Fraassen, Bas (1995): "Belief and the Problem of Ulysses and the Sirens", *Philosophical Studies* 77 (1), 7 – 37.