Suppositions, Revisions and Decisions

Daniel Y Elstein and J Robert G Williams

September 2, 2014

Abstract

Supposition can seem philosophically perplexing. It comes in two different forms, and it is not immediately obvious what one of these forms (supposition in a "counterfactual" mode) is good for. This paper relates supposition-as-counterfactual to rationality—arguing that it should be at the heart of causal decision theory and hypothesis testing. We explain how familiar formulations of decision theory can be understood in terms of this core formulation. Conditional chance, we argue, is a norm for counterfactual supposition, and we explore how exactly it should be formulated. Another tradition formulates causal decision theory through counterfactual conditionals. We argue that if such formulations line up with those in terms of supposition and chances, their principal utility will be to enable deliberation and communication of practical reasons. This gives us an attractive unified package—of practical rationality articulated in terms of counterfactual supposition, with the latter normed by chance and communicated via conditionals.

Contents

Suppose Nixon had pressed the nuclear button—back in the 70's, alone in his office. The Russians, no doubt, retaliate to the US aggression. Other nations are dragged into the conflict. Devastation ensues. Suppose, on the other hand, that Nixon *did* in fact press the nuclear button—back in the 70's, alone in his office. The conclusion would have to be that the failsafes in the system kicked in and the president called it off before the weapons were fired. After all, we know that nuclear war (thank goodness) never occurred. Thus, suppositional reasoning comes in two forms. For in the example above, one and the same content—that Nixon pressed the button—is the object of supposition; but the consequences our suppositional reasoning draws out are very different.

This paper discusses the point and purpose of the first of these modes of supposition, and its relationship to other notions—in particular, to (conditional) chance and to counterfactual conditionals. Part I introduces the basic framework of supposition; its role in rational decision making; and the conditions under which it is correct. Part II examines some complex issues about the relation between chance, supposition and time. Part III takes a wider view on the relation between suppositions so characterized and counterfactual conditionals.

I The framework

I.1 Two kinds of supposition

Supposition can be done in various styles. The Nixon cases above illustrate a case where two sensible-seeming modes of suppositional reasoning come apart. This distinction is systematic and widespread. It seems intimately linked to the distinction between indicative and counterfactual conditionals. Adams' famous pair illustrates this: "If Oswald hadn't shot Kennedy, someone else would have" and "If Oswald didn't shoot Kennedy, someone else did" (Adams, 1970). Again, in considering the former conditional, we suppose in one way that Oswald didn't shoot Kennedy; in considering the latter conditional, we suppose the same proposition in a different way. The two modes of supposition, we take it, are familiar phenomena; it is useful to have labels for them. Call the former 'supposition as counterfactual', 'C-supposition' for short, and the latter kind 'supposition as actual' or 'A-supposition'.[1]

The two kinds of supposition can deliver different verdicts about whether a given proposition is true-under-a-supposition: under the A-supposition that Oswald didn't shoot Kennedy we take it to be settled that someone other than Oswald shot Kennedy; whereas under the C-supposition that Oswald didn't shoot Kennedy we take it to be unsettled whether someone other than Oswald shot Kennedy. The proximate explanation of the difference is that under the A-supposition the fact that someone shot Kennedy is still available to introduce as a premise in reasoning, whereas under the C-supposition that fact is in some way screened off. An account of the different kinds of supposition will involve giving a deeper explanation of this phenomenon.

What is A-supposed is treated like new evidence, whereas what is C-supposed is not. When we A-suppose that Oswald didn't shoot Kennedy, we consider what happens when that proposition is added to our stock of evidence. We have considerable evidence that Kennedy was shot which in no way depends on it being Oswald who shot him; so even if we updated on new evidence that Oswald did not shoot Kennedy, we would still believe that someone shot Kennedy, because our other evidence for this would be undefeated.[2] The difference with C-supposition seems to be that it concerns worldly rather than evidential connections. To decide on whether someone else would have shot Kennedy under the C-supposition that Oswald did not, we need to know what the setup of the world was prior to Oswald's shooting of Kennedy. For instance, was there someone else planning to shoot Kennedy? In that case running the clock forward from that prior point but somehow omitting Oswald's shooting still results in someone shooting Kennedy, leading us to assent to the counterfactual. When we C-suppose something we do not believe, the question is what (imagined) reconfiguration of the world would best accommodate the truth of what is C-supposed. A-supposition depends on evidential connections between beliefs; C-supposition on causal and other worldly connections between events.

Whatever detailed account of the difference between the two kinds of supposition we end up giving, it is at least clear that there are two kinds. But what's the point and purpose of these men-

[1] The names here recall Jackson's distinction between A-intensions and C-intensions (cf. Jackson 2000). This is intentional, though having two 'dimensions' of modality does not give one a story about the details of two kinds of supposition.

[2] This is a contingent feature of the case. Compare: "If the Bolsheviks did not shoot the Romanovs, someone else did". This is false if our only evidence that anyone shot the Romanovs is evidence that the Bolsheviks did. If we updated on new evidence that the Bolsheviks did not shoot the Romanovs, we might well come to believe that nobody shot them.

tal states? For A-supposition the story seems straightforward: we want to decide how the world is, but there is much that we don't know; to cover all the 'live' options we will need to consider how the world is on various suppositions which cover the epistemic possibilities. That seems useful along multiple dimensions. But what's the point of having in addition, C-supposition? Why would we want to think about the consequences of hypotheses we antecedently know to be false?

Perhaps one point to C-supposing known-false hypotheses, is that we thereby gain modal information: information about how things could have or would have been. But that just pushes the challenge back a step. Why, after all, should we care about the merely possible? Whilst it is plausible that there is a link between C-supposition and modal knowledge, perhaps we would do better by understanding the importance of the latter in terms of the importance of the former.[3]

There are two functions served by C-supposition whose point is not merely modal. The first is epistemic. We may C-suppose that some hypothesis holds, work out what follows, and compare the results with what we already know about the world, rejecting the hypothesis when they do not match. The hypothesis must be C-supposed, not A-supposed, in order for this procedure to count as a genuine test. If we've just observed an explosion, then (since the knowledge isn't "screened off" in any way) on almost any sensible A-supposition, the explosion occurs. This means that we cannot test a hypothesis about the past by A-supposing it and seeing whether the explosion occurs on that A-supposition. On the other hand, it might well be that under the C-supposition that chemical X were present, the explosion wouldn't have occurred; while under the C-supposition that chemical Y were present it would have. C-supposition, and its "screening off" of items of knowledge, thus seems to serve a genuine epistemic need.[4] This rationale extends to C-supposing things we take to be false. Even when we know something to be false, we may have an interest in bolstering our knowledge with additional justification; or in revisiting our reasons to check they are cogent; or in going through our arguments to convince others they should think as we do.

The second function of C-supposition is within practical rationality. The next section re-

[3]Williamson (2007) argues for an epistemology of modality via knowledge of counterfactuals, and takes the latter to work via C-supposition. Divers and Elstein (2012) argue that at least a necessary condition for believing that P is absolutely necessary is being prepared to introduce P as a premise under any C-supposition.

[4]This point is made by Stalnaker (1999, 70-1) and emphasised by Williamson (2007, 138).

views the basic contrast between evidential and causal decision theories which, following Joyce (1999), we present as based on A- and C-supposition respectively. But familiar presentations of causal decision theory bring in further resources—causal backgrounding, counterfactuals or chance, for example. We will argue that the right way to see this as making a separable claim about the right way to C-suppose.

I.2 Two kinds of decision theory

Consider Newcomb's problem:[5] Predictor (who from past experience seems faultless in predicting your behaviour) presents you with two boxes. Box 1 is transparent and contains £100. Box 2 is opaque, and contains either £1000 or nothing; it contains £1000 if Predictor predicted that you would open only Box 2, and it contains nothing if Predictor predicted that you would open both boxes. You can keep whatever money is in a box that you open. Should you two-box (open both boxes), or one-box (open just Box 2)?

The argument for one-boxing: given Predictor's past success you think it overwhelmingly likely that Predictor will have correctly predicted your choice, and so if you two-box it is almost certain that Box 2 is empty and you will end up with £100, whereas if you one-box it is almost certain that Box 2 contains £1000, and so you will end up with £1000. Only one-boxers get rich, and getting rich is better than not, so better to one-box.

The argument for two-boxing: however good Predictor is, the prediction is already made and the contents of Box 2 are settled, so by one-boxing you are simply passing up the £100 in Box 1, which makes you worse off whatever is in Box 2. One-boxers get rich, but that doesn't make them rational; it is just that Predictor predicted that some people would irrationally one-box and decided to reward their irrationality (or so say two-boxers).[6]

Two familiar versions of decision theory can be seen as embodying these opposing arguments. Evidential decision theory (EDT) endorses one-boxing: your choice is *evidence* for what is in Box 2, and since you want to maximise expected value your choice must be sensitive to the way this evidence influences rational expectations. If you one-box, your one-boxing is conclusive evidence that Box 2 contains £1000, so the expected value (V) of one-boxing is

[5] See (Nozick 1969). The problem was invented by William Newcomb.
[6] See (Lewis 1981a).

£1000 (or near enough, if there is some tiny chance that Predictor was wrong). If you two-box, your two-boxing is evidence that Box 2 contains nothing, so the expected value of two-boxing is £100 (or near enough). Since the expected value of one-boxing is higher than the expected value of two-boxing, you should one-box.[7]

Causal decision theory (CDT) endorses two-boxing: although one-boxing and two-boxing provide different evidence for what is in Box 2, they do not *influence* what is in Box 2, and only the differential effects of your choices can make a difference to their relative expected value (U). Since the contents of Box 2 are fixed at the moment of choice, the expected value of two-boxing is simply £100 higher than the expected value of one-boxing, whatever the Predictor predicted (two-boxing dominates one-boxing).

EDT and CDT both speak of maximising expected value; it is just that they have different conceptions of expected value. We call the EDT version of expected value V, and CDT version U. Both V and U are functions from a utility function u defined over all outcomes (complete possible worlds), and a (subjective) probability function P defined over all propositions. The formulation of V is well-understood: the expected value V of the proposition A (e.g. that you perform a particular action) is the sum of the utilities of the outcomes (assigned by u), with each utility multiplied by the probability of its outcome on the A-supposition that A is performed (assigned by P). We take $S_1, S_2, \ldots.$ to be a partition of states. These are propositions that, at least in conjunction with A itself, completely specify one of the outcomes. And we write the probability of a proposition p on the A-supposition of A as $P_A(p)$. So then we can characterize V in line with EDT as follows:

$$V(A) = \sum_i P_A(S_i)u(A \wedge S_i)$$

This is not the most familiar way of writing the definition of V. It is more usual to talk in terms of the conditional probability of each state, $P(S_i|A)$, rather than the A-suppositional probability $P_A(S_i)$. But the difference is bridged by identifying A-suppositional probability with conditional probability. That is natural given the Bayesian idea that updating on A (adjusting one's credences on learning exactly A) goes by conditionalization, when we think of supposing

[7]For an influential development of EDT, see (Jeffrey 1965).

A as a kind of simulated updating.[8]

One can formulate CDT by direct analogy with EDT. In each case, we assess the value of actions by figuring out the likely outcome of the action on the supposition it is performed. The value is the weighted average of the utilities of the outcomes, with the weights in each case given by the suppositional probability of that being the outcome. In the case of EDT the perspective is evidential, and the weights are given by A-suppositional probabilities, $P_A(S_i)$. For CDT, the weights should reflect the probability of S_i given the worldly impact of A—the extent to which A being true influences whether S_i is true.

Unsurprisingly, we suggest that the weights are given by C-suppositional probabilities of an outcome under the supposition that an act is performed. Writing the probability of a proposition p on the C-supposition of A as $P^A(p)$, we can then characterize U as follows:

$$U(A) = \sum_i P^A(S_i)u(A \wedge S_i)$$

Note that on this presentation the only difference between V and U, and hence between EDT and CDT, is that the former uses A-suppositional probability, and the latter uses C-suppositional probability.

Just as with the formulation we gave of V, this is not the most familiar presentation of CDT (though it is essentially the "common core" formulation that Joyce (1999) recommends we focus on). Moreover, there's no widespread and uncontroversial formal treatment of the key notion of C-supposition, playing the role that conditional probabilities do for A-supposition. The two issues are linked. The various reformulations of CDT that are popular in the literature can be viewed as diverging on what account they give of C-supposition. So construed, the reformulations do not conflict with the formulation just given; rather, they attempt to make its core notion formally tractable by connecting it to other notions—counterfactuals, relevant causal factors, chance, and so forth. It is the task of the following sections to explore these connections, and the accompanying claim that the various formulations are equivalent. Let us emphasize, however, one advantage of using a C-supposition formulation: it connects up naturally with the psychological reality of decision making, which plausibly does involve considering what would

[8]Though there is the difference that when I suppose that p, it can be true on that supposition that I don't believe that p; whereas when I update on p I am bound to then believe that I believe that p.

happen supposing one were to perform this or that action.

The upshot of the discussion so far is that by thinking about C-supposition and CDT together we can avoid problems which arise when we consider each on its own. If, as the majority of philosophers believe, CDT norms instrumental rationality, then on our preferred formulation of CDT it turns out that C-supposition is heavily involved in the correct decision theory for instrumental rationality. This is a decisive answer to the concern that C-supposition lacks a function to explain why we care about it.

Within CDT we face the problems of deciding which of the competing formulations to settle on, and of providing a clear story about the psychological resources which reasoning in accordance with CDT will draw on. Thinking of CDT in terms of C-supposition can help here, because C-supposition is a psychological phenomenon independent of CDT; its nature contrasts with A-supposition in the same way that CDT contrasts with EDT; and taking C-supposition to be central to CDT, we shall argue, provides a way of evaluating alternative formulations.

I.3 Formulations of causal decision theory

Practical rationality involves supposing you take an action, and considering the desirability of various outcomes and their likelihood under that supposition. CDT involves C-supposing you take those actions. So if we want to know how to make decisions properly, we need to know how to C-suppose properly.

There are constraints on (theoretically) rational C-supposition. If Sian thinks that she would clear a tall building, on the C-supposition that she attempts to leap over it, she's very likely irrational. For that C-suppositional mental state to be rational, Sian would need very special background beliefs—the belief that her muscles are so strong, on the gravitational environment so weak, that her attempt would have a high chance of success. If she has those background beliefs, her overall belief state may be rational even if she's sadly mistaken about her musculature or the gravitational field (due perhaps to misleading evidence). If she lacks those background beliefs, her mental state is internally messed up.

C-suppositions are rationally constrained by one's wider mental state (compare: A-suppositions and conditional degrees of belief, as characterized by the ratio formula). If we can determine the proper value of $P^A(p)$, for arbitrary suppositions A and propositions p, in terms of categorical degrees of belief P, then the notion will become formally tractable. And by substituting in this characterization into our generic formulation of CDT above, we'll get a new formulation of CDT—one appropriate at least under the idealizing assumption that our decision-making is theoretically ideal.

Indeed (following Joyce) we can back-engineer candidate rational constraints on C-supposition from the varieties of CDT that exist in the literature. Here are three:

1. **Conditional chance**

   Sian is aware that the objective chance of B given A is 0.8. Plausibly in such circumstances Sian's C-suppositional probability for B on A should be 0.8 too. That suggests a broader hypothesis: one's C-suppositional probabilities are accurate when they match the corresponding conditional chances. Further, when one is uncertain what those chances are, rational C-suppositional probabilities should match one's expectation of conditional chance.[9] That is:

   $$P^A(B) = \sum_x x \cdot P(Ch(B|A) = x)$$

   Requiring C-suppositional probability to match expected conditional chance seems the natural analogue of the requirement that (unconditional) credence match expectations of chance: the "Principal principle".[10]

[9] A quick argument for the move to expectations given the accuracy-claim just made. Suppose that we score C-suppositional beliefs for 'accuracy' in world w, by how distant $P^A(B)$ is from $Ch(B|A)$. By a theorem of De Finetti, appropriately interpreted, ones C-suppositional beliefs will be guaranteed to be 'accuracy dominated' unless they are expectations of conditional chance. Similar arguments can be formulated for the parallel moves from claims about correctness to claims about expectations for the rivals below.

[10] The analogy suggests a puzzle case. Suppose you take yourself to have foreknowledge of the results of a chance process—what you take to be a reliable Oracle has told you that the fair coin spinning in the air will land heads. What should be your C-suppositional probability of winning, if you bet on it landing heads? Going by your expectation of the chances, this should be 0.5. But if you trust the Oracle, you might think it should be near 1. Perhaps, then, the norm on C-supposition needs a rider that makes it inapplicable in cases where one has 'inadmissible information about the result of relevant chances processes.

However, recall that one of our paradigm applications of C-supposition was in hypothetical reasoning about the causes of past events. We need to be able to evaluate the plausibility of setup S on the basis that, C-supposing S to be in place, observation O is unlikely; yet O came about. That the chance of O given S is low is excellent backing for such claims. But our knowledge that O came about would exactly be inadmissible information about the outcomes of the relevant chance processes. So we think that 'inadmissibility exceptions are the wrong way to go. Its a very nice question about what we then say about rational decision-making where we take ourselves to have oracular information about the results of chance processes (if it is ever rational to believe that)—but well

2. **Counterfactuals**

Gibbard and Harper (1978) influentially formulated CDT in terms of the probabilities of counterfactuals. In the current framework, we can understand their approach as saying that the accurate C-suppositional probability of B on A will match the truth value of the counteractual $A \mathrel{\Box\!\!\rightarrow} B$, and accordingly rational C-suppositional probabilities will match the expected truth value of the corresponding counterfactual, which is to say that they will match the credence one invests in that counterfactual:[11]

$$P^A(B) = P(A \mathrel{\Box\!\!\rightarrow} B)$$

One could try to justify this approach by the simple thought that to ask whether B is true on the C-supposition that A is simply to ask whether the counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ is true; and so to ask whether B has a certain probability on the C-supposition that A is to ask whether the counterfactual has that probability. If Sian has 0.6 probability in the counterfactual, $A \mathrel{\Box\!\!\rightarrow} B$, doesn't she have to have 0.6 probability in B on the C-supposition that A?

3. **Counterfactual chance**

Conditional chances of B on A are facts about the relative weight that the actual chances put on B-possibilities, within all A-possibilities. An alternative idea (Lewis 1981b) is to relate C-supposition to facts about the way that the chance of B would have been, had A occurred. Suppose that were Sian to flip the coin she's holding, the objective chance of that coin landing heads would have been 0.7. Then—very plausibly—Sian should have 0.7 probability in the coin landing heads under the C-supposition that she flips it. That suggests another broad hypothesis: one's C-suppositional probabilities are accurate when they match true counterfactuals about chance. Further, when one is uncertain over

leave that complex of issues for another day.

[11]Within a standard worlds-formulation of the conditional, the above equation is only tenable with broadly Stalnakerian treatment of the conditional, on which the underlying selection function can be used to characterize directly an 'imaging probability that C-suppositional probabilities can match. Joyce shows that generalized imaging functions for less demanding treatments of conditionals will satisfy an inequality. The C-suppositional probability of B on A is bounded below by ones credence in $A \mathrel{\Box\!\!\rightarrow} B$, and bounded above by ones credence in $A \mathrel{\Box\!\!\rightarrow} \neg B$. Our discussion of the counterfactual norm below will generalize to these imaging treatments.

which such counterfactual is true, rational C-suppositional probabilities should match one's expectation of the counterfactual chance. That is:[12]

$$P^A(B) = \sum_x x \cdot P(A \mathrel{\Box\!\!\rightarrow} Ch(B) = x)$$

Each proposal says something about what the accurate C-suppositional belief to hold is, given facts about chances and counterfactuals that obtain in the actual world. Each extends this to a story about rational constraints on categorical beliefs about chance/counterfactuals and C-suppositional beliefs. Each gives us the resources to say why Sian is irrational if she C-supposes that she would clear a tall building were she to attempt to leap it if she has *normal* views about the corresponding chances/counterfactuals concerning success. On the other hand, If she has *unusual* beliefs about the gravitational field or her musculature, then her views on chances and counterfactuals might make her C-suppositional belief perfectly rationally held. That is again a good thing. To be sure, in that circumstances something has gone wrong, but the wrongness is the wrongness of inaccuracy not of irrationality.

Still further formulations of CDT are available (see Joyce 1999 for a systematic survey), but we will not explore the relation between the notions discussed here and e.g. causal background partitions. It should be clear how to extend the kind of discussion to follow: one starts by assuming that the crucial difference between EDT and CDT is really a difference between A-supposition and C-supposition; given that, rival formulations of CDT can be viewed as further theses about C-supposition, to be evaluated in the light of the package-deal that emerges at the end of this paper.

At this stage, it's not obvious whether the three norms under discussion diverge in their recommendations for C-supposition. Later we'll examine that closely. Conceptually, however, there's an obvious difference. The conditional chance norm appeals only to chances themselves; the counterfactual norm appeals only to counterfactual conditionals; the counterfactual chance norm presupposes both an antecedent grip on chances and on counterfactuals. That will be significant when we come later to consider the relation between C-suppositions and beliefs in

[12] Just as before, if we accept that accurate C-suppositional probabilities match the counterfactual chances, then an accuracy-domination argument supports the conclusion that under conditions of uncertainty the rational constraint is to match expectations of counterfactual chance.

counterfactuals.

In this paper, we will be defending the conditional chance norm, and arguing that counterfactuals have an expressive rather than normative role. In Part II we assume rather than arguing for the conditional chance norm—the aim being to elaborate exactly how it is to be understood and applied to CDT and hypothesis testing, with particular attention to the question of which temporal perspectives are the relevant ones to takes chances from. With this more fleshed-out proposal in mind, we are able in Part III to see how the conditional chance norm interacts with counterfactuals, and why it is superior to alternative proposals of counterfactual or counterfactual chance norms on C-supposition.

II Chance, Time and Supposition

II.1 Temporal perspectives and backtracking

In Part II, we will be assuming that expected (conditional) chance norms C-supposition. But chances change over time. At 3 o'clock, there was a chance that Billy would be called away to do the shopping before 4, preventing him from finishing watching the film. At 4, having finished the film, there's no longer any chance of this—time has rolled on, making some events more likely and eliminating certain prospects altogether. Suppose Billy was deciding whether to agree to go to his film club that evening. He very much wants to be there—but he'd be humiliated at turning up without having watched the film. If Billy is practically rational, in making this decision he will C-suppose agreeing (/not agreeing) and evaluate the likelihood and value of the outcome. C-supposition is normed by the expected chances. But which chances?

One way that this issue arises is with the conditional chance norm articulated earlier.[13] In the light of the time-relativity of chance, we need to subscript the appeals to chance on pain of ambiguity:

[13] As we shall explain later, the counterfactual chance norm seems to have a built-in privileged time to take chances, just after the time that the counterfactual supposition/antecedent is made true. This means that the counterfactual chance norm is in one respect at an advantage: it deals neatly with at least the simple cases of date-range supposition which we discuss below. But as we explain later, the inflexibility of the counterfactual chance norm ends up being a disadvantage overall: it makes it impossible to use it to give a unified account of the role of C-supposition in hypothesis testing as well as decision theory.

$$P^A(B) = \sum_x x \cdot P(Ch_t(B|A) = x)$$

And again we ask: for a given A and B, what is the relevant "temporal perspective" from which to C-suppose?

We shouldn't immediately assume that there must be an informative context-free answer to this question. One interesting proposal would be that there are many legitimate ways to C-suppose, one for each "temporal perspective" reported by the t. Call that the variantist perspective. The opposite view, on which for a given A and B there is a privileged temporal perspective, we call invariantist.

Such differences wouldn't matter if all temporal perspective delivered the same norms on C-supposition. But they do not. Consider the following scenario. Rhodri is trapped in a tower. The present pattern of guards make it very unlikely that any escape attempt would succeed. That both makes the present conditional chance of success, given an attempt at escape, very low, and also makes it very unlikely he'd make the attempt in the first place.

Add in the following feature. Rhodri could last night have put in place a plan to drug all the guards (he didn't in fact do this). Escape attempts with drugged guards very likely succeed. At 7pm yesterday, the likeliest scenario where Rhodri makes an escape attempt the next day was by screwing up his courage to drug the guards that evening, then shinning down the drainpipe the next day—a timid child, this wasn't very likely, but much more likely than his making an unprepared escape attempt.

What this means is that within all scenarios where Rhodri makes an escape attempt, yesterday's chances place highest weight on scenarios where the guards are drugged beforehand. In most of those, he succeeds. So *yesterday's* conditional chance of success, given Rhodri attempts to escape today, is high. This differs from *today's* conditional chances because today's conditional chances bake in the fact that he did not drug the guards yesterday evening.

The upshot of this is that if it is yesterday's conditional chances that are relevant to the C-supposition that Rhodri attempt an escape, then our norm predicts a high probability of success; and if it is today's conditional chances that are relevant, then our norm predicts a low probability of success. The temporal perspective can make a crucial difference. (The difference is closely connected to the distinction between standard and backtracking readings of counterfactual conditionals (Lewis, 1979). The latter distinctively involve some interpolated reasoning about the most likely way of the scenario being realized. In the case at hand, that corresponds to factoring in the fact that Rhodri is more likely to attempt an escape with preparation rather than without. The 'standard' reading treats more of the actual history as fixed.)

The differences that choices of temporal perspectives induce also matter for decision making. Consider again Newcomb cases. Let t be the time just prior to Predictor making her prediction and putting money in the boxes. Let t* be the (later) time at which the agent must choose whether to 1-box or 2-box. The agent, we shall suppose, has no information about what Predictor has done. The recommendations based on t and t* temporal perspectives are very different.

Consider $Ch_{t^*}(\text{Empty}|\text{2-box})$—the conditional chance at t* of Box 2 being empty given that the agent 2-boxes. In a situation where Predictor has predicted 2-boxing, this conditional chance is approximately 1, but in a situation where Predictor has predicted 1-boxing, the conditional chance is approximately 0. But crucially, since Predictor has already put money in the boxes, the agent knows that $Ch_{t^*}(\text{Empty}|\text{2-box}) = Ch_{t^*}(\text{Empty}|\text{1-box})$.[14] And so the agent also knows that $Ch_{t^*}(1100|\text{2-box}) = Ch_{t^*}(1000|\text{1-box})$, and $Ch_{t^*}(100|\text{2-box}) = Ch_{t^*}(0|\text{1-box})$. This means that the expected value (U) of 2-boxing is known to be greater than that of 1-boxing, no matter how likely it is that Predictor predicted 2-boxing.

Notice, however, that this story crucially relies on using t*-chances. If we switch focus to the time before the predictor makes the decision, t, there is no reason to think that $Ch_t(\text{Empty}|\text{2-box}) = Ch_t(\text{Empty}|\text{1-box})$. Indeed the Predictor's reliability suggests that $Ch_t(\text{Empty}|\text{2-box})$ is close to 1, and $Ch_t(\text{Empty}|\text{1-box})$ is close to 0. The setup thus ensures that $Ch_t(100|\text{2-box})$ and $Ch_t(1000|\text{1-box})$ are both near 1. Being aware of this, if we use t rather than t* as our temporal perspective for C-supposing, we should expect much less money supposing we 2-box than we get supposing we 1-box. And using those C-suppositions within CDT would produce the advice

[14] A potential confusion on this point concerns the idea of expectations of conditional chance. We imagine the following complaint: "Surely the agent's expectation of $Ch_{t^*}(\text{Empty}|\text{2-box})$ should be higher than the expectation of $Ch_{t^*}(\text{Empty}|\text{1-box})$ because, given Predictors accuracy, the box is more likely to be empty if the agent 1-boxes." The crucial reason why this is mistaken is that the relevant expectations of conditional chance are not based on supposing the relevant condition; they are based on whatever information the agent has about the actual situation. If the agent fairly confidently expects to 2-box, then the agents expectation of both $Ch_{t^*}(\text{Empty}|\text{2-box})$ and $Ch_{t^*}(\text{Empty}|\text{1-box})$ should be high.

to 1-box. This may be a surprising result: it is normal to talk as though the dispute between 1-boxers and 2-boxers is between EDT and CDT. But if it is right that the real difference between EDT and CDT is between A- and C-supposition, then CDT only sides with 2-boxers on the assumption that C-supposition is normed by the chances at t* (the moment before decision) rather than chances at an earlier time such as t.

What is going on here? The reliability of the predictor means that, at time t, the action that the decision maker performs and the prediction made (and thus distribution of money in the boxes) are not probabilistically independent. The probabilistic dependence is however, "screened off" once the prediction is made. For example, if the predictor plumps for 1-boxing, we get the result that at t* there's a low chance that 2-boxing occurs (that's required if the predictor is reliable) but the die is cast—there's no longer anything the predictor can do to tie the money in the boxes to what in fact occurs. So even if the later chances of 2-boxing is low, the chance of getting 1100 given one 2-boxes is at the later time close to 1. As chances evolve over time, conditional chances become disentangled.

II.2 A free choice of temporal perspective?

The moral of the preceding section is that temporal perspectives can make a big difference to what a norm on C-supposition predicts. One prima facie attractive package would be the following. First, one has a free choice in choosing the temporal perspective relative to which one C-supposes. Second, when C-supposing for a certain external purpose (e.g. for the purpose of making a decision) only some temporal perspectives are relevant.

An attraction of the first response is that it frees us from the responsibility of giving an abstract, context-free story about what the temporal perspective is relevant to a given C-supposition. It also allows us to predict and explain the presence of both "standard" and "backtracking" styles of C-supposition (just as with counterfactual conditionals, the backtracking way of engaging with Rhodri's escape attempt seems a legitimate way to engage with the supposition, and so we should seek an account with the flexibility to accommodate it). Finally, when we eventually do want to say that some temporal perspectives are better than others, we can say that it is better relative to some specific purpose.

Thus, it's consistent with all this flexibility in C-supposing that we maintain that for the special case of C-supposing in order to decide what to do at t, the relevant temporal perspective is just before the moment the decision is implemented.

But we earlier identified another role for C-supposing, in hypothesis testing. What the proper choice of temporal perspective is for those cases needs to be considered independently. An interesting case involves hypothesis testing involving common causes. Imagine that you are in the same room as someone who claims to hear loud thunder. You suspect her of lying. Why? Because if she had heard thunder, you would have heard it too. (Imagine that you take her to be far more likely in general to lie than to hallucinate.) That is, to test the hypothesis that your companion heard thunder, you C-suppose that she did, and reason under that C-supposition that you would hear thunder too; you do not hear thunder, so you reject the hypothesis (or at least doubt it).

Why should you believe that you would hear thunder, on the C-supposition that your companion hears thunder? On our account this would be vindicated by the conditional chance $Ch(\text{you hear thunder}|\text{she hears thunder})$ being high. Temporal perspectives now matter. Imagine that the temporal perspective was 'just before the supposition came about', to parallel the decision theory case. From that temporal perspective (the time of the C-supposition), then the conditional chance of you hearing thunder given that she does will not be high, because at that time it is settled that there was no thunder to hear (there was no lightning a few seconds earlier). Given that there was no thunder to hear, the only way that your companion hears thunder is by hallucinating it, and there is no reason to think that you will share her hallucination. So at t, Ch(you hear thunder—she hears thunder) is low. It is only if we shift to t*, an earlier time at which lightning might have taken place that you each might hear at t, that the relevant chance is high.

It is fairly common to test a hypothesis by looking for the effects of the underlying cause of the hypothesis holding, rather than the direct effects of the hypothesis itself. The use of C-suppositional reasoning in such cases, which aims to discover the truth by tracking worldly connections, must track back to the causal source of those connections in order to do its work successfully. So whilst the purposes of CDT set the relevant time as that of decision, the purposes of hypothesis testing may set the relevant time as at least as early as that of the latest potential causal link between the hypothesis and the test (which is earlier than both if they would have a common cause). A certain time-asymmetry matters to decision-making, but is irrelevant in finding out about the world. We decide what to do in light of the differential effects of our actions, given the set-up of the world at the time of action, so it is only the downstream causal connections that matter. In the epistemic case there is no reason to ignore the causal connections upstream of a hypothesis being tested. If C-suppositions themselves don't bake in a temporal perspective, we are free to cut our cloth to suit the contours of the application under consideration.

II.3 Why variantism is insufficient

For all its theoretical attractions, this will not do. The difficulties arise when we consider suppositions that are not about the the obtaining of one particular dated event. There are plenty of such suppositions, e.g. of generic claims (suppose kangaroos had no tails) or temporally unspecific happenings (suppose Roosevelt had been assassinated). But we will concentrate on date range suppositions. These concern the obtaining of a localized event, but only specify when it happens within a certain range. Thus, we can consider what would have happened under the supposition that the bookshelf collapsed last week; that Rhodri tried to escape one day last week; or that Sian played the Newcomb game one day last week (with Predictor placing money into boxes the evening before).

The temporal variantist will see a huge variety of options when faced with a date-range scenarios. But they struggle to predict what seems like it has a good claim to be a 'standard' way of supposing. Let's take the Rhodri case earlier. Rhodri was in the tower all last week, made no escape attempt, and every day the guards watched attentively. The way of supposing we search for is characterized by two features:

Supposing that Rhodri had made the escape attempt one day last week:

1. It might have been on Monday

2. Whichever day he attempted it, he likely failed.

Aspect (2) of this engagement with the date-range supposition is simply the analogue of disambiguating to a non-backtracking reading. To be sure, we could engage with the supposition in a backtracking way, and work backwards to Rhodri having made preparations for escape by drugging the guard. But equally we can work forward, noting that given the guards were in fact constantly attentive, any escape attempt would be doomed.

What temporal perspective would secure this result? The problem is that if we set the temporal perspective early in the week—Sunday evening, for example—then the most likely scenario on which Rhodri makes an escape attempt will be later on in the week, having made adequate preparations. By setting the temporal perspective early in the date-range, we generate backtracking predictions for realizations of the supposition later in the date-range.

One solution to this would be to fix upon a late perspective. If we choose the temporal perspective on Friday lunchtime, there's no time left for Rhodri to make any preparations, and so we rule out backtracking readings. But this then violates constraint (1)—the temporal perspective bakes in a 'late branching' scenario excluding the Monday escape.

For all the variety that variantism affords, it can't give us everything we want. The trouble is that it specifies a single temporal perspective exogenously. Intuitively, for a date-range supposition like this, we want the flexibility to adopt different temporal perspectives depending on the different ways that the supposition could be realized. Thus, on the realization on which Rhodri makes the escape attempt on Monday, it is Monday's temporal perspective that is relevant, on the realization on which Rhodri makes the escape on Tuesday, it is Tuesdays, and so forth. Another way of putting this is that there should a reading of the date-range supposition on which the following argument pattern rationally constrains the suppositional attitudes, with each of the premise-attitudes rationally held:

- Zero probability of success under C-supposition that Rhodri attempts escape on Monday
- Zero probability of success under C-supp that Rhodri attempts escape on Tuesday
- Zero probability of success under C-supp that Rhodri attempts escape on Wednesday
- Zero probability of success under C-supp that Rhodri attempts escape on Thursday
- Zero probability of success under C-supp that Rhodri attempts escape on Friday

Therefore:

- Zero probability of success under C-supposition that Rhodri attempts escape between Monday and Friday.

The date-specific suppositions involve varying choices of temporal perspective; and if the conclusion involved a single temporal perspective, it would be unconnected from four of the five premises.

The problematic presupposition of the discussion thus far is that we read the time-indexes in the chance norms on supposition rigidly—that t in the formulation of the conditional-chance norm (repeated here) is a rigid designator, to be filled in with an exogeneous choice of temporal perspective:

$$P^A(B) = \sum_x x \cdot P(Ch_t(B|A) = x)$$

But we've seen this is the wrong reaction. We need to allow that which time is relevant varies. Here is one way this could happen in the context of one of the other takes on C-supposition introduced earlier: the counterfactual chance norm. Let us view 't' as a non-rigid designator, something like 'the time just after A was made-true',[15] and formulate it thus:

$$P^A(B) = \sum_x x \cdot P(A \mathbin{\Box\!\!\rightarrow} Ch_t(B) = x)$$

Thus, if A is the date-range proposition that Rhodri made an escape attempt last week, the denotation of t will vary depending on which Rhodri-escape-world we consider. If the escape was made Monday, then t denotes Monday; if on Tuesday, then t denotes Tuesday, and so forth. This, of course, is exactly the variation of temporal perspectives we need in order to generate the forward-tracking reading of the date-range supposition that leaves open all relevant possibilities, just as it should.

A simple non-rigid designator reading of t would work for the counterfactual chance norm because the t occurs within the scope of a counterfactual operator. Making an analogous tweak for the conditional chance norm on C-supposition is more difficult. To be sure, if t is construed as non-rigid, then an agent uncertain about which world is actual will then be considering dif-

[15]In the counterfactual chance formulations, the natural choice of time-index is just after A is brought about, rather than just before—the reasons for this are given in the next section.

ferent temporal perspectives in the different worlds they are open to. But in something like the Rhodri tower case, there is no relevant uncertainty about the way the actual world is. We know that Rhodri was stuck in the tower all week.

If we want to make this work, we need to ensure that we tie together a temporal perspective and ways of realizing the supposition. One idea is the following. In the case of date-range suppositions, the conditional chances that are relevant are not conditionalized on the supposed content alone. Rather, there are some specific realizations of the supposed content. In supposing that Rhodri attempted escape last week, what is relevant are chances conditional on a Monday-escape, Tuesday-escape, Wednesday-escape etc. If we let g(p) denote an arbitrary realization of p, then we can let T(g(p)) denote the time just before g(p) occurs.[16] The chance norm then takes the form:

$$P^A(B) = \sum_x x \cdot P(Ch_{T(g(A))}(B|g(A)) = x)$$

To what extent is this a move away from variantism? Insofar as variantism presupposes rigid designation of times in the chance norm, it is unacceptable for date-range C-suppositions. But a defeat for variantism here is not a victory for invariantism. The dispute between variantism and invariantism focuses on date-specific C-suppositions: the variantist wants to allow for both forward-tracking and backtracking uses, whereas the invariantist does not. The issue in this section concerns the connection between date-range and date-specific suppositions, and the central

[16] In the case as described we can see escape attempts on particular days as representing specific realizations, the chances of escape are the same whichever day the attempt is made, and the chances of an escape attempt being made are also the same for the different days. This is why it makes sense to reason about an arbitrary realization in the way described. But what happens when matters are more complex: if it is not so clear how to divide the date-range into specific realizations, or the different realizations result in different chances of success, or the realizations are not equally likely? This is a hard problem, and it may be that in such cases it is indeterminate what the relevant conditional chance is, and so indeterminate what the C-suppositional probability should be. Such a result might be problematic if date-range C-suppositional reasoning has a role in decision-making via CDT. (Suppose, for example, that you have to decide whether to adopt a plan to escape on some unspecified day between one and two weeks in the future.) Perhaps what is needed is an average of the chances conditional on the different realizations, weighted by their relative likelihood (this can be understood as the unconditional chance of that realization occurring, with the relevant time to take chances being relative to the date of the potential realization e.g. just before). That would seem to harmonise well with the spirit of CDT. Such a solution could be understood as telling us which conditional chance is relevant to a date-range supposition; alternatively it could be seen as a recipe for getting a sensible answer out of CDT when it is indeterminate what the right conditional chance is. Either way, we are less confident that this weighting procedure has as much claim to being simple common sense as does the move to low chance conditional on a date-range supposition from low chance conditional on an arbitrary realization of the supposition (when things are equal between the realizations). The resolution of this issue seems sufficiently complex and controversial that it is beyond the scope of this paper; we have indicated, however, that the defender of the conditional chance norm does seem to have resources for dealing with even the less straightforward date-range C-suppositions.

datum is simply that we take the chances conditional on a date-range supposition to depend on the chances conditional on the date-specific suppositions which would count as relevant realizations of the date-range supposition. This issue should in principle be separable from that between variantists and invariantists. The variantist can tweak the interpretation of the above formula as follows: instead of saying that T(g(A)) denotes the time just before g(A) occurs, say that it denotes the time relevant to assessing the C-suppositional probability of B on g(A). When (according to the variantist) our interests require forward-tracking, the relevant time will indeed be just before g(A) occurs; when backtracking is required it will be some earlier time.

Let us step back a bit, and consider the significance of the discussion we've just been having. A *general* theory of rational C-supposition must have something to say about date-range suppositions. The generality is important to its epistemic function, since the hypotheses that we test against observations may well take a date-range form. One might wonder, however, whether someone interested in C-supposition solely for the purposes of CDT needs to worry about it. One might think that actions that a decision theory evaluates are always dated events (movements of the body, paradigmatically). For the narrow purposes of CDT, the thought would be that we could get away with a story about norms on C-supposition that restricts itself to these special cases. But this is probably false. Some of our rational decision making consists in choosing between relatively unspecific *plans*. The plan at issue might well be something like making an escape attempt next week. We want to bring the resources of our best theory of decision to bear on what plan to adopt. But then we exactly want to consider what is likely to eventuate under an unspecific, date-range supposition. The CDT-ist cannot afford to ignore the issues we have been grappling with here.[17]

What have we learned in this Part? First, that the temporal perspective matters a great deal what a chance norm on C-supposition will predict. Second, that there are attractions to a liberal perspective in which no particular temporal perspective is baked in to C-supposition of a particular content as such. But finally, that the assumption that there is a single, rigid temporal perspective relevant to counterfactual supposition is both unargued for and produces

[17]One way to view a decision theory, is to see it as articulating rational constraints between belief, supposition and *desire*, with the connection to action being given by the principle that one chooses the most desirable option for action available. If that is correct, then the characterization of a decision theory needs to be wide enough to cover any content one might assign a level of desirability—and that will surely include date-range propositions.

the wrong predictions, eliminating the non-rigid readings that are crucial to securing salient forward-tracking readings of date-range suppositions in particular. Exactly how to formalize the relevant kind of non-rigidity is a tricky issue for the conditional chance norm, but it's clear that some such story is needed.

We remain friends of the conditional chance norm, appropriately refined and qualified. But others might see in the complexities we have just identified reasons to reject in in favour of other candidate norms on C-supposition. In the next few sections, we argue that this would be a very bad idea.

III Suppositions, conditionals and decisions

Part I introduced C-supposition and some candidate norms. Part II investigated one norm on C-supposition, based on conditional chances. In Part III we will be reintroducing counterfactual conditionals. Section 7, immediately below, identifies two different roles that counterfactual conditionals can play in our account of rational decision and update—they could provide norms on C-supposition, or they could be vehicles for expressing such C-suppositions. In section 8, we argue that counterfactuals alone can't play the normative role. In section 9, we extend this to the counterfactual chance norms of Part I. Section 10, meanwhile, raises a challenge to the idea that counterfactuals can play the expressive role. Section 11 takes on that challenge and identifies several ways in which counterfactuals can play the expressive role.

III.1 Practical reasoning and conditionals

Instrumental practical rationality addresses the following question: given what you take the world to be like (your doxastic state) and what you want to bring about (the desirability of outcomes by your lights), what actions should you take? This should be distinguished from questions of theoretical rationality—in particular, whether your doxastic state is itself as it should be. One can be non-ideal in either dimension: either by having a bad doxastic state, or by acting or intending to act in ways that are not supported by your beliefs and desires.

The normative construal of the connection between C-supposition and counterfactuals and/or chance contrasts with a reductive construal of the same equations, on which (e.g.) C-suppositional probability just is a mixture of unconditional probabilities in propositions whose content is about something else. That would make it impossible for the equations to be violated; and impossible for those who don't have the concept of chance/counterfactuals to have C-suppositional attitudes at all. Both claims seem unwelcome. C-supposing is a familiar psychological state, and one can engage in it without thinking about counterfactuals/chance in the right way (whatever that is). Of course, if one is considering appropriate ideally rational agents, then ex hypothesi such norms are met, and one can 'reformulate' CDT in terms of chances/counterfactuals, rather than appealing to C-supposition directly. But such reformulations may be misleading guides to the non-ideal case—and one benefit of being clear about the distinction between (theoretical) norms on C-supposition and (practical) norms on decision-making is it allows us informative categorization of lapses from perfect rationality.

Our proposal above was that the C-suppositional formulation is the right formulation of what it takes to be practically rational. Quite independently of that normative connection between beliefs, desires (/actions/intentions) and C-suppositions, one can ask: which C-suppositions are the right ones to have, and what patterns of categorical beliefs and C-suppositional beliefs are rational to hold together? That question broaches issues of theoretical rather than practical rationality, and the conditional chance norm gives our answer to this.

Acting in theoretically and practically rational ways is one thing; explaining to oneself or others why one acted as one did is another. The difference is reflected in a further distinction: between intentional and deliberative rationality. An intentional explanation of action points out those mental states that make the action intelligible. On the view under consideration, this would be a combination of basic desires for outcomes, and C-suppositional beliefs. An agent might act in intentionally rational ways, while lacking the ability to represent to themselves or others why they acted as they did. They may have C-suppositional beliefs, whilst lacking the concept of C-supposition itself (which, after all, is a fairly sophisticated notion). But if they had the concept of C-supposition—or some other concept deploying which would convey information about C-supposition beliefs—then they would be able not only to act rationality, but also articulate why they acted rationally. That's a considerable advance. With such a representation of reasons for action, they can then reflectively evaluate whether those reasons were good ones both practically (whether they relate to the action taken in the right way) and theoretically (whether the beliefs were the right ones to have).

It's a striking fact that practical reasoning explanations often appeal to counterfactual conditionals. "Why did Sian run towards the wall and jump, injuring herself?" "Because she thought that were she to attempt to jump the building, she would impress her friends by jumping right over it." The face-value reading of this practical-reasoning explanation of Sian's action appeals to her belief in a counterfactual conditional truth. It doesn't mention chances or C-suppositions at all. An overall theory of practical rationality needs to account for the fact that it's sensible to use counterfactual conditionals in deliberative contexts such as these (or to recommend a revision of extant practice).

One way to account for the role of counterfactual conditionals in deliberative rationality is to hold that they already have a role in intentional rationality. The formulation of CDT in terms of conditional chances does not do this; but the other two proposals introduced in I.3 do. In Sian's case, the view would be that what makes her action practically rational is her high probability in success given the C-supposition that she attempted to jump the house; but that C-supposition in turn is rationally connected to a high confidence in a corresponding counterfactual conditional. The counterfactual conditional here is either about the chance of success or the fact of success; the gap between intentional theory of decision and explicit practical reasoning explanations is either narrowed or eliminated altogether.

This makes a norm on C-supposition that features counterfactual conditionals *prima facie* attractive. But notice that it does place a particular burden on one's theory of counterfactual conditionals. For them to provide an independent *norm* on C-supposition, we need a grip on counterfactual conditionals that is *independent* of C-supposition itself. Lewis's (1979) account of counterfactual truth-conditions is an example of the sort of thing we are after (at least for the case of 'standard' readings of counterfactuals he targets). Plugging in this Lewisian account, the right way to C-suppose would be spelled out in terms of counterfactual truths; counterfactual truths would reduce to facts about comparative similarity of worlds; and similarity of worlds reduces to facts about laws and duplication. But we'll see that this requirement for an independent grip on counterfactuals is a strategic weakness in these proposals.

Building counterfactual truths into our story about the intentional explanation of actions via a counterfactual-featuring norm on C-supposition isn't the only way of accounting for their presence in practical reasoning explanations. A different strategy is endorse an equation such as $P^A(B) = P(A \mathbin{\Box\!\!\rightarrow} B)$, not on the right-to-left reading on which counterfactual truths norm C-supposition; but on a left-to-right reading on which C-suppositions are the basic phenomenon, and the right credence to have in a counterfactual is just the corresponding C-supposition. C-suppositions norm belief in counterfactuals rather than vice versa, on this view. Rather than an independent grip on counterfactuals, we'd need an independent grip on correct C-supposition—a salient candidate would be the conditional chance norm from section I.3. The upshot would be that while the story about the intentional rationality of action is counterfactual-free, counterfactuals have a prominent role to play in deliberative rationality—reporting a belief in a counterfactual is a canonical means of expressing our C-suppositional beliefs. This sort of 'suppositional' approach to conditionals is very familiar in the context of A-suppositions and indicatives, where it has been defended by Adams, Edgington and Bennett among others (cf. Bennett 2003 for an extended survey). It is far less prominent in the literature on counterfactuals (though some of the authors just mentioned have advocated a unified approach, and Skyrms 2004 explicitly advocates the relationship between chance and counterfactuals just sketched).

We think that the role of counterfactuals in deliberative rationality is a datum that must be accounted for, but we favour the second explanatory strategy. In the next two sections we give our reasons for favouring conditional chance norms over counterfactual (/chance) norms. After that, we consider whether the role of counterfactuals in deliberation can be squared with this emerging picture.

III.2 Against counterfactual norms

This section argues that counterfactual norms on C-supposition are highly problematic. This will support the view that the ubiquity of counterfactuals in practical reasoning explanations is a phenomenon of deliberative rationality—the expression of our reasons—and not by their having a role in intentional explanations.

We emphasized above that if counterfactuals are to norm C-supposition, we must have some independent grip on how counterfactuals work. The strategy of this section is to identify an assumption that is common in one tradition for giving an independent grip on counterfactuals.

We'll come back at the end to consider what more general dialectical morals can be drawn from these considerations.

The key assumption that we work with is the following: Some (not all!) possible but low-chance outcomes conditional on antecedent A can be "counterfactually excluded" by A (at least in worlds where A is in fact false). That is, we have true instances of $A \mathrel{\Box\!\!\to} \neg B$, where conditionally on A obtaining, B still has a chance of obtaining. Lewis's treatment of chance and counterfactuals is a case in point. On that view, worlds that contained 'quasi-miracles' (surprising low-probability events, like a ball quantum-tunnelling through the floor after being dropped) are ipso facto pushed further out in the similarity-ordering than worlds containing no quasi-miracles. Williams' (2008) treatment of these cases has the same effect, though the criterion under which quantum-tunnelling scenarios are pushed further out is that they are atypical regions, relative to the chancy laws of the world. In these cases, there's some (tiny) finite chance that the atypical/quasi-miraculous quantum-tunnelling occurs, given that the ball is dropped. But in worlds where the ball is in fact not dropped, the theory predicts that were the ball to be dropped, the quantum tunnelling would not occur.

These claims about the ordering of worlds have a rationale. Ordinary counterfactual assertions often presuppose the counterfactual exclusion of the kind of phenomenon Lewis would describe as quasi-miraculous. Rhodri says that if he had dropped the fragile glass on the stone floor, it'd break and he'd have to sweep it up. That seems mundane and obviously true—and yet the consequent entails that the glass broke rather than quantum tunnelled. The usual logic of counterfactuals ensures that the counterfactual exclusion of Lewisian quasi-miracles follows from these mundane counterfactuals.

Quasi-miracles needn't be based in quantum mechanics. The kind of freak occurrences that statistical mechanical probabilities gives low but non-zero probability of obtaining (see Elga 2001) would undermine the truth of mundane counterfactuals. For present purposes, we'll use a low-key placeholder for these quasi-miraculous events, one suggested by Hawthorne 2005 and discussed by Williams: atypical sequences of infinite or long-finite coin flips. So what we'll be assuming is that the following is determinately true (in non-billion-flip worlds): "if I were to flip this fair coin a billion times, it would not come up heads every time". It would, however, be straightforward to adapt the cases to any other instance of this kind of counterfactual exclusion of the atypical/quasi-miraculous.

We use counterfactual exclusion to make trouble for counterfactually-normed CDT. Suppose you are faced with...

The Devil's game
You have to decide to PLAY or NOT PLAY.

- NOT-PLAY: delivers status quo (utility 0).
- PLAY: a fair coin is flipped N times;
  - If N heads, HELL (-100 utility);
  - otherwise, SMARTY (1 utility).

The generic CDT recipe for evaluating whether you should take this bet is to average HELL and SMARTY utilities, weighted by the C-suppositional probabilities that the respective outcome eventuates, given that you play.
Assuming C-suppositional probabilities go by conditional chance, this gives the sensible verdict that the chances of hell have to be sufficiently small to justify running the risk—i.e. N has to be sufficiently large. Further, in variants of the Devil's game where the disutility of hell is worse (the fires are turned up, the stay is longer) the critical point gets further away.
But now suppose that C-supposition goes by counterfactuals. In not-PLAY worlds, ALL-HEADS/HELL outcomes are counterfactually excluded, ex hypothesi. So the informed will have conditional credence zero in the counterfactual PLAY$\Box\!\!\rightarrow$ ALL-HEADS, given they don't in fact play. This tells us immediately that one who was antecedently certain they wouldn't end up playing (and is well-informed of counterfactual truth-conditions), will give no weight at all to the HELL outcome. It is always rational for them to PLAY, no matter what the objective chances may be, or how bad HELL is. But it equally has implications for those antecedently uncertain about how they will act. PLAY$\wedge$ALL-HEADS worlds may make the PLAY$\Box\!\!\rightarrow$ ALL-HEADS counterfactual true (so called "strong centering" would ensure this), so rational agents give some weight to HELL, if they give some credence to playing. But the HELL outcome is still underweighted. For the agent's credence in the counterfactual is a mixture of their credences in it conditionally on PLAYING and on NOT-PLAYING respectively. We know their credence in the counterfactual, conditional on the latter, is zero.

But their credence in PLAY$\Box\!\!\rightarrow$ ALL-HEADS, given PLAY and strong centering, is simply their credence in the conjunction PLAY $\wedge$ ALL-HEADS. In general, the disutility of HELL will factor into counterfactual-based CDT in the devil's game, but will be underweighted, to an extent proportional to the agents' credence that they will play. For an agent who knows and sets their credences in the current chances, the weight of the disutility of HELL by the lights of chance-based CDT will be $Ch(\text{ALL-HEADS}|\text{PLAY})$ while the weight according to counterfactual-based CDT will at most $Ch(\text{ALL-HEADS}\wedge\text{PLAY})$.

This result is just wrong. One way to bring this out is to consider cases parallel to the Devil's game, but where HELL eventuates when some slightly lower chance outcome obtains, compared to the original setup. Perhaps this is the conjunction of some specific random sequence of coin flips (this has equal chance to the all-heads outcome but won't count as atypical/quasi-miraculous), together with the obtaining of some event with arbitrarily high $(1-\varepsilon)$ probability. In a forced choice between the original Devil's game and this variant, it would be crazy to go for the one that had the higher chance of HELL. It would be crazy, that is, to favour the original Devil's game over this parallel one. Yet that is what counterfactually-based CDT recommends, because the informed will not assign zero credence to the counterfactual "If I were play the new game, HELL would result" conditional on not playing.[18]

The argument boils down to this. Mundane counterfactuals exclude certain atypical events. So such atypical (but positive-chance) events would not happen, given mundane antecedents. But then if we build C-supposition around counterfactuals so construed, the positive-chance atypical outcomes are zero-weighted (or at least underweighted) when it comes to considering the possible outcomes of actions. But that can easily be massaged into cases where the recipe gives obviously terrible advice.

One reaction is that we must avoid a theory of counterfactuals that excludes atypical events. That is a reasonable reaction, and the minimal moral from the considerations raised above is

[18]You might try a more sophisticated form of counterfactual-based CDT, where we appeal to imaging probabilities rather than probabilities of counterfactuals. But this doesnt make a difference. Joyce shows that the imaging probability of B on A will be bounded above by the probability of the counterfactual $A \Box\!\!\rightarrow \neg B$. But in the case at hand, we can argue that in $\neg$ PLAY worlds the probability of PLAY$\Box\!\!\rightarrow$ ALLHEADS is 1. And so by the bound principle we get that the imaging probability, just like the probability of the counterfactual, is zero. So we can rerun the argument.

simply that one cannot responsibly endorse a counterfactual-based CDT while giving a blank cheque to the theory of counterfactuals that underlies it. Specifically, one needs to tackle the puzzles that motivate counterfactual exclusion of the atypical/quasi-miraculous, and pinpoint how to simultaneously respond to those puzzles and get reasonable results in a counterfactually-based CDT. What we suspect happens is that theorists tempted by this direction leave counterfactuals as a working primitive, and simply assume that in the cases that matter it will not give silly results. But that strategy leaves open that pre-theoretic opinions about counterfactuals are relying on the guidance of C-supposition (perhaps normed by conditional chances), rather than identifying a body of propositions on which we have independent grip, that can then be used to norm C-supposition. We'll see some backing for this suspicion in a couple of sections' time.

We have a second objection to the counterfactual norm which arises from the use of C-suppositional reasoning in hypothesis testing. As we have explained, hypothesis testing typically proceeds by C-supposing the hypothesis, working out what follows from that supposition, and comparing with our evidence to see whether the hypothesis is confirmed or disconfirmed. When the probability of B given H is higher than the probability of B given $\neg H$, finding out that B goes some way to confirming H, and finding out that $\neg B$ goes some way to disconfirming H. If the counterfactual norm is correct, then the relevant sense of 'probability of B given H' is understood in terms of the probabilities of counterfactuals (rather than conditional chances and the like).

The crucial thing to note is that our use of this method of hypothesis testing in no way depends on the temporal order of H and B: it is fine if B is earlier than H. This works for the conditional chance approach, so long as we take chances early enough. Let B be that an atom of the sample decays within a particular time interval, and let H be that the cat dies within a particular, later interval. Suppose that things are hooked up so that the decay of an atom of the sample will almost certainly bring about the cat's death, the cat is unlikely to die at the relevant time if no atom decays, and, given the relevant half-life, decay is unlikely. So at some time t before B, $Ch_t(B)$ is small and $Ch_t(B|H)$ is very high. So if we rely on conditional chances at t, then the hypothesis that the cat died is disconfirmed by evidence that no atom decayed. Can the approach that appeals to counterfactuals rather than chances as the norm on C-supposition replicate this success?

First of all, notice that in order to get the correct result, the counterfactual norm would have to read the relevant conditional as a backtracker, since the consequent concerns a time prior to the antecedent (a forward tracking reading would, disastrously, identify our expectation of $H \mathrel{\Box\!\!\rightarrow} B$ with our expectation of B). We emphasized earlier that the counterfactual-norm on C-supposition requires we have an independent grip on the the counterfactuals; we see here that it is not enough to point to an account of forward-tracking counterfactuals (such as Lewis 1979); we also have to have an independent grip on backtracking counterfactuals. These tend to be in short supply.

But second, notice what is happening. The conditional chance norm (if the temporal perspective is appropriately picked) gets the right answer straight away. We're exploring whether there's any reading of counterfactuals that replicates the success. But that means that *at best*, what one will be able to show is that our credence in counterfactual will coincide with the expected conditional chance, so that the two norms speak with one voice in their recommendations. The best case scenario is one where the counterfactual norm tells us nothing that we couldn't have gotten already out of the conditional chance norm. But things may be much worse than that. We'll see in Section III.10 that there's a threat that we cannot maintain an equation between the probability of the counterfactual and the expected conditional chance—that will mean that the conditional chance and counterfactual norms would diverge in their recommendations, and (we assert) it's conditional chances rather than counterfactuals that would then be epistemically significant. The sense that there *must* be some such reading of the counterfactual on which credences in counterfactuals are fit to play this role in hypothesis testing, we suggest, reflects the fact that our expectations of the truth values of counterfactuals are ordinarily guided by our beliefs about conditional chance.

III.3 Against counterfactual chance norms

We have just given two reasons to be unhappy with a counterfactual norm on C-supposition. Can a norm invoking counterfactuals about chance do better?

Let's briefly recap the distinction between conditional chances and counterfactual chances.

The conditional chance of B given A, $Ch(B|A)$, is simply defined as $\frac{Ch(B \cap A)}{Ch(A)}$, so it is a ratio of unconditional chances. If we know all the unconditional chances (at t), then the conditional chances (at t) can simply be read off from them. To go by conditional chances is still to go by actual world chances. In contrast, the counterfactual chance of B on A is the (unconditional) chance that B would have had were A to obtain: $A \mathbin{\Box\!\!\rightarrow} Ch(B)$. In Part II, we explored what sort of temporal subscripts we need to add to the chances involved in the conditional chance norm; parallel issues arise for the counterfactual chance norm. We'll start by reviewing the parallel issues here.

Lewis 1981b argued that counterfactual chances should be indexed to the time *just after* the moment at which A is realized. That tends to work well at least for paradigm applications within CDT. It's worth noting the problems that would arise with other choices for t—in particular if t is chosen to to be *before* the occurrence of A. The trouble is that the t-chances of B may very well remain low, even on the counterfactual assumption of A, no matter how strong the worldly dependence between A and B is. Suppose A and B are both unlikely at t (0.1, say), but A-type events invariably cause B-type events. There's no reason to think that the counterfactual t-chance of B occurring is high, were A to occur. (Some special cases illustrate this: e.g. if A actually does occur, and strong centering holds, then the counterfactual t-chance of B is just its actual t-chance, and so is 0.1.)[19] So no matter what else we might add to Lewis's suggestions about the choice of time-index for the counterfactual chances that putatively norm C-supposition, we should reject any that are temporally prior to the supposition itself. This little lemma will be helpful in our two arguments below.

We will focus on two objections to the counterfactual chance norm that parallel those we just made to the counterfactual norm. The first is that when the counterfactual chance and conditional chance norms disagree in cases relevant to CDT, it is the conditional chance norm which seems to produce the more plausible results. The second is that the conditional chance norm can generalise to the use of C-supposition in hypothesis testing in a way that the counterfactual chance norm (like the counterfactual norm) cannot, and this means that adopting the conditional chance norm gives us a more unified account of C-suppositional reasoning.

[19] A limiting case is one where $B = A$. Then surely on any decent notion of C-supposition, the C-suppositional probability of A on A is 1, and indeed for any t later than the occurrence of A, $A \mathbin{\Box\!\!\rightarrow} ch_t(A) = 1$, but if A was unlikely, we could easily have $A \mathbin{\Box\!\!\rightarrow} ch_t(A) = 0.1$.

Regarding the first objection, the counterfactual chance advocate can satisfyingly respond to The Devil's game. Take the basic case that caused trouble for a counterfactual norm: we suppose the actual world is a ¬PLAY one. Were you to play the Devil's game, the objective chances of HELL would be small but non-zero. To be sure, it's also true on these theories that were you to play, the coin would not land heads, and so you would not get HELL (that's what caused the trouble before). But those counterfactuals are not what norm C-supposition on the counterfactual chance approach—it's the counterfactuals about chance that matter, and we've seen as yet no reason to doubt that they will deliver sensible results, and so sensible recommendations for action when paired with CDT.

But we may be able to get the problem back. The recipe will be to massage the exclusion of the atypical so as to make certain *chance distributions* counterfactually excluded. Let us take the PLAY option in the earlier example to be partitioned according to whether a atypical/quasi-miraculous event Q obtains or not. So the two states to consider are PLAY$\wedge Q$; PLAY$\wedge\neg Q$. In this altered game, the devil gives you the SMARTY in PLAY$\wedge\neg Q$ worlds, and gives you HELL in PLAY$\wedge Q$ worlds. Were you to PLAY, it would be the case that PLAY$\wedge\neg Q$, by the exclusion of the atypical/quasi-miraculous, and so you would get SMARTY. But in the present context, the crucial question is what the counterfactual *chance* of Q would be, if you were to PLAY. If we can choose Q so that is is zero, we can rerun the original arguments, since counterfactual chances will underweight the prospect of HELL.

Consider the following setup. Let Q be a *way* of realizing PLAY, so that $Q\wedge$PLAY or $\neg Q\wedge$PLAY is realized simultaneously with PLAY (perhaps PLAY involves you waving a white flag, and Q has you waving it vigorously so it collides, but then quantum tunnels, through the wall next to you; or perhaps it involves you conveying your acceptance while very quickly flipping nervously a sufficiently large number of coins, with them all landing heads). What's the counterfactual chance of of Q, in PLAY worlds? Recall our lemma above: the only sensible time-indices for counterfactual chance are *after* PLAY. Since Q is resolved simultaneously with PLAY, the chances at that time are either 1 or 0, depending on whether or not Q obtains in the world in question. So just as PLAY counterfactually excludes Q, it counterfactually ensures that Q has chance zero (at the relevant time).[20]

[20]If Q concerned an event earlier than PLAY, in the portion of history that is plausibly 'held fixed' in the closest

The above argument shows that given an influential view about counterfactuals and quasi-miracles, the counterfactual chance norm will have incorrect implications for action in certain cases. The caveat of the previous section applies again: the case built up depends on one particular theory of counterfactuals, and so one could escape the argument by rejecting that kind of theory. But the burden is on the proponent of counterfactual chances to tell us what the alternative theory is that both deals with the puzzles that motivate the counterfactual exclusion theses, and returns sensible verdicts. The move to counterfactual chances, rather than bare counterfactuals, makes no progress on this front.

Our second objection to the counterfactual chance norm returns to the use of C-supposition in backtracking hypothesis testing that was deployed in the last section as an argument against the counterfactual norm. Recall the point that when working out whether some potential observation, B, confirms or disconfirms a hypothesis, H, it is fine for B to be temporally prior to H, and the conditional chance norm handles this smoothly. The argument against the counterfactual chance norm proceeds much as before. Think of things in terms of a dilemma: either we take counterfactual chances relative to a time before B or after B. The first horn takes t to be earlier than B. But then, since B is earlier than H, t ends up being earlier than H. By our earlier lemma that is unacceptable (it will make the counterfactual chances of B match the actual chance of B at that time, uninfluenced by the obtaining of H).

What then, of the second horn? If t is any time after B, we know that $Ch_t(B)$ will be either 1 or 0, matching the truth value of B. Writing $|B|$ for the truth value of B, 1 or 0, the expectation of the x such that $H \mathrel{\Box\!\!\rightarrow} Ch_t(B) = x$ will exactly match the expectation of the y such that $H \mathrel{\Box\!\!\rightarrow} |B| = y$. So on this horn of the dilemma, what norms C-supposition in this instance of hypothesis testing would be the probability of the truth value of the counterfactual conditional: $H \mathrel{\Box\!\!\rightarrow} B$. The second horn of the dilemma then, is that the counterfactual chance norm collapses into the counterfactual norm in such cases, and we have already seen why that is unsatisfactory: (i) we need the norm to select backtracking counterfactuals as relevant; (ii) it seems that the counterfactual norm (and in turn the counterfactual chance norm) is acceptable only insofar as it approximates the conditional chance norm (and we'll shortly see some

counterfactual world, then it would not necessarily be counterfactually excluded—exact match of particular fact trumps minimizing quasi-miracles, by Lewisian lights. But our Q has been choosen to be part of the history that must change for PLAY to be brought about, so there's no escape from the conclusions here.

reasons for worrying that it can't even do that). So it turns out that the best case for the counterfactual chance norm is that it somehow reiterates rather than replaces the conditional chance norm.

We have raised two independent objections to the simple counterfactual norm on C-supposition, which both extend to the counterfactual chance norm. Assuming that there is a unified norm for C-supposing in practical and epistemic uses, we only need one of them to work to undercut the proposals. There is good reason to reject counterfactual-involving norms on CDT.

III.4 Can we have it all?

Let's reflect on where we've got to. The idea that conditional chances norm C-supposition is, we think, overwhelmingly plausible (as plausible as the Principal Principle). We have urged that expected chance can't replace C-supposition in our basic formulation of decision theory, but that in the ideal limit the formulations will coincide. And we've identified two potential roles for counterfactual conditionals in this setup—normative and deliberative.

If the world were nice, it'd turn out that we could have it all—conditional chances would norm C-supposition (coinciding for ideal agents). Likewise, for ideal agents C-supposition would coincide (and perhaps explain) credences in counterfactuals/imaging probabilities—thus allowing us to explain the common practice of expressing our reasons for action in terms of counterfactual conditionals.

Unfortunately, there is a conflict between the expected chance formulation of CDT, and the counterfactual/imaging formulation. In order for them to agree, we need to have the following (for ideal agents):

$$P^A(B) = \sum_x x \cdot P(Ch(B|A) = x)$$

(If with Joyce we favour an imaging formulation, we will replace the above with an formulation whereby the expected chance is an upper bound on the probability of the counterfactual). But such an identity (or bound) is highly problematic. In the special case where an ideal agent has full information about what the chances are, she will match her credences to the chance of the counterfactual, and the right hand side will simplify to a single conditional chance. We will have:

$$Ch(A \,\Box\!\!\rightarrow B) = Ch(B|A)$$

(Again, in Joyce's version, with a bound rather than identity) The form of such a principle should be familiar—it equates a probability (in this case, chance) of a conditional, with the corresponding conditional probability (chance). And we are then neck-deep in the territory of the "triviality" and "impossibility" results kicked off by Lewis (1976). To be more specific: under widely accepted assumptions, there just is no two-place conditional operator satisfying this equation. (The adaption of the Lewis impossibility result to such formulations is discussed in Williams (2012), which also includes an argument that the bounds formulation is no better).[21] Unless these results can be resisted, then the search for harmony between chance-normed and counterfactual/imaging formulations of decision theory is quixotic.
The tension might have been anticipated. After all, consider how the parallel issues would play out with EDT. There, the overwhelming consensus is that the A-suppositional probabilities figuring in EDT calculations match the conditional probabilities (at least where the latter are defined). But to deliberate and communicate the EDT-rationality of a given action, one needs a device that allows one to express high conditional probability. If the probability of an indicative conditional was the conditional probability (or even provided a bound for it) we could link EDT practical rationality with ordinary (indicative-conditional invoking) practical reasoning explanations. But notoriously, the required equation is subject to the barrage of impossibility results. What we have noted here is that we do not escape this issue by shifting to CDT, chance and counterfactuals.

III.5 Taking stock

Where does this leave us? As already indicated, it is difficult to deny that conditional chance norms C-suppositional probability, and so is crucial both to CDT and hypothesis testing. This is indeed the package we advocate. But this leaves a residual puzzle: our reports of our

[21] Further triviality results for counterfacutals can be found in Leitgeb (2012) and Briggs (ms). Schwarz (ms) contains very instructive comparison and discussion.

C-suppositional reasoning, both in ordinary practical reasoning explanations and in hypothesis testing, are often phrased in terms of counterfactual conditionals. How can we account for this fact? We will focus on the practical case first, and then argue that a similar account applies in the epistemic case as well.

One option is to finesse the impossibility results—to argue that the counterfactual formulation is not in fact a rival to the others. Bradley (2012) is a recent and very sophisticated attempt to tackle exactly this issue specifically focusing on the relationship between counterfactuals and conditional chance. Beyond this, a model for strategies can be found in the various ways in which the impossibility results for indicative conditionals have been treated. For example, one might hold that counterfactual conditionals do not express propositions in the first place, but rather directly express C-suppositional probabilities. This is the analogue of the Adams-Edgington position on indicative conditionals, on which they have no truth-conditions, but instead express A-suppositional probabilities. Giving up truth-conditions for counterfactuals is likely to be highly unpopular in many quarters (though see Skryms 1994, Edgington 1997), but it has the considerable local virtue of (a) allowing a counterfactual formulation of CDT that—ideally—coincides with the chance and C-supposition varieties; and (b) in virtue of this, making sense of use of counterfactuals in practical reasoning explanations. Of course, there's no plausible normative role for counterfactuals, given the direction of explanation involved. But that's not a worry, if chance already plays the normative role. Other familiar strategies for the indicative case might be transferred and tried out—see van Fraassen (1976), McGee (1989) and Jackson (1987) for some famous strategies.

A second option is to accept the divergence, but make the case that credences in counterfactuals are good enough approximations for practical purposes of deliberation and communication. If ceteris paribus, the probability of counterfactuals near enough matches expected conditional chance, diverging markedly only in recherch cases that we rarely if ever face, then perhaps we can make sense of the fallback strategy. We would urge anyone tempted by this sort of line to try to make precise the sense of "approximation" and "other things being equal" that occur in it—and presumably we could not fully evaluate its prospects without a specific account of counterfactuals to work with.

We do have a working suggestion to offer, based on some ideas (again for the indicative case) put forward in recent years by Frank Jackson. Start with the observation that, whatever else we might later say, people's intuitions about counterfactuals do seem to track their expectations about conditional chance. Indeed, cases where the probability of counterfactuals comes apart (relative to certain theories) from known conditional chance have been offered as objections to those theories. So it seems pretty likely that we do, as a matter of fact, conform our credences in counterfactuals to expectations of conditional chance. This may well be seen as widespread error if the controversial "reconciliation" strategies don't work out. Maybe the best deserver for the role of "counterfactual conditional" will not vindicate our pre-theoretic attitudes (Jackson thinks this is so for the indicative analogue: the best deserver for being the indicative conditional, he urges, is the material conditional.)

What is important is again the non-ideal case. If there is a widespread practice of treating counterfactual credence as coinciding or providing a bound for C-suppositional probabilities, then this is something that agents can exploit. If I assert a counterfactual conditional, I expect you to think that my C-suppositional probabilities are high. And so I can use the counterfactual conditional as a device for conveying such information (or treat them as such in my internal deliberation). Maybe fully informed agents would be rightly queasy about this practice, but the data we had to explain was exactly concerned with ordinary, non-ideal agents.

So our suggestion is the following: acknowledge the intuitive force of identifying C-suppositional probabilities with the credence in counterfactuals, use this bare data to explain observed practice, and keep an open mind about whether or not intuitions are vindicated by the best theory of counterfactual conditionals.

But we also mentioned an epistemic rationale. In wrapping up, let us point out that the two rationales are linked, and the debate will in many ways run parallel.

The epistemic role for C-supposition, recall, was that it allowed us to test a hypothesis by C-supposing H, and seeing whether what followed matched our evidence (A-supposition is ill-suited to this task, we noted). The quantitative setting is again natural: C-supposing that the coin has a weight distribution that we know leads to a 9/10 bias should lead us to expect to a certain degree to see it land heads when flipped; and we assess the success of the hypothesis by seeing how well our C-suppositional credences match observations. It's clear that conditional chance should be a norm for C-supposition for these purposes—otherwise outcomes that are entirely in line with the chances would partially disconfirm H (taking care, as discussed above, to focus on the right temporal perspective from which to take chances).

Notice that hypothesis testing and CDT are, on this account, based on the same underlying use of C-suppositions as a way of applying beliefs about chance. What the two cases have in common is that they use C-supposition as a tool for applying our knowledge of (chancy) worldly connections to something undecided: whether to believe a hypothesis, whether to perform an action. Since in each case a proposition is up for grabs (in one case epistemically, in the other practically) we cannot simply take it as a premise for ordinary, non-suppositional reasoning, and must instead take that proposition as a C-supposition and reason in a way disciplined by our views on the relevant conditional chances.

The parallels continue in the relationship between C-supposition, chance and counterfactuals. Our degree of belief that we'd hear loud music, under the C-supposition that Jimmy were home is high. But we don't hear music. On this basis, we conclude with some confidence that he's out on the town again. Our ordinary engagement with this sort of thinking simply needn't involve articulate thought of chances. If credences in counterfactuals aligned with C-supposition and hence with expected conditional chance, then we'd have a package that explained naturally the connection between "hypothesis-test" results of this kind and our ordinary practice of communicating and deliberating such reasoning using counterfactuals. Lewisian "impossibility" results make this package suspect, of course.

If we end up saying that counterfactual probability and conditional chance diverge, then just as in the case of CDT it seems like the correct epistemic norms will tell us to test hypotheses by attending to the latter rather than the former. In that case our use of counterfactuals in hypothesis testing would need to be explained, since we do in fact use counterfactuals to report our C-suppositional reasoning: "If Jimmy were home, we'd hear loud music; we don't hear any music; so he isn't at home". Again we could say that counterfactual talk successfully communicates C-suppositional probability and thus conditional chance, even if strictly speaking its semantics is quite different.

IV Conclusion

Let us go back to the beginning. C-suppositional belief can seem, at first blush, mysterious. What's the point of supposing something one knows to be false? What's the point of screening off (”ignoring”) one's evidence, when supposing? A-supposition, nicely described by conditional probabilities, is a kind of attitude it's easy to see the function for. Its counterfactual cousin can seem the poor relation—a device, perhaps, for discovering truths about the merely possible (and that can then seen an optional extra, unconnected to our ordinary epistemic and practical concerns). This paper maintains that C-supposition, far from being an idle wheel, is the core of epistemic practice and rational agency, as captured in CDT.

Here is the picture of C-supposition we would like readers to take away. C-supposition is a dyadic proposition attitude, not to be reductively identified with beliefs about counterfactuals, chances or anything else. Beliefs about conditional chances do play a normative role with respect to C-suppositions. Just as chances are relative to a time, C-suppositions are relative to a temporal perspective (care is needed in formulating what this comes to, since we must allow such temporal perspectives to be “non-rigid”). Which temporal perspective should be adopted depends on the purposes for which we are C-supposing. Relevant purposes include hypothesis-testing and decision making. Counterfactuals don't figure in this story about the intentional rationality of action or update; but they loom large in the corresponding stories about deliberative rationality, wherein we represent to ourselves our reasons for acting or updating. The tool may be perfectly suited to do its job (if a theory such as Skyrms' or Bradley's works out) or may be imperfect but good-enough (as on the Jackson-inspired proposal above).[22] Since the role of counterfactuals in this context is to allow us to reflect on our rationality, rather than as part of the basic story about rationality itself, we need not worry that any imperfections end up committing us to absurd conclusions about epistemic or practical rationality.

[22] The authors of this paper are inclined to go different ways on the question of which is the most promising direction.

Bibliography

Adams, E. W. 1970: ‘Subjunctive and Indicative Conditionals’. Foundations of Language, 6, pp. 89-94.
Anderson, Alan Ross. (1951) ‘A Note on Subjunctive and Counterfactual Conditionals’, Analysis, 12(2): 35-8.
Bennett, Jonathan (2003) ‘A philosophical guide to conditionals’. OUP.
Bradley, Richard (2012) Multidimensional possible-world semantics for conditionals’ Philosophical Review, 121 (4).
Briggs, Rachael (Unpublished): ‘Two Interpretations of the Ramsey Test’
Divers, John and Elstein, Daniel Y (2012) ‘Manifesting belief in absolute necessity’ Philosophical Studies 158 (1):109-130.
Edgington, Dorothy. (1997) Truth, objectivity, counterfactuals and Gibbard’, Mind, 106 (421): 107-116.
Elga, Adam (2001) ‘Statistical Mechanics and the asymmetry of counterfactual dependence’. Philosophy of Science.
Gibbard, Allan and Harper, William (1978). Counterfactuals and Two Kinds of Expected Utility. In A. Hooker, J. J. Leach and E. F. McClennen (eds.), Foundations and Applications of Decision Theory. D. Reidel. 125-162.
Jackson, Frank. (1987) Conditionals (Oxford: Blackwell).
Jackson, Frank. (2000) From Metaphysics to Ethics (Oxford: OUP).
Jeffrey, Richard. (1965) The Logic of Decision (Chicago: University of Chicago Press).
Joyce, James. (1999) The Foundations of Causal Decision Theory (Cambridge: Cambridge University Press).
Hannes Leitgeb (2012): ‘A Probabilistic Semantics for Counterfactuals. Part A’. The Review of Symbolic Logic, 5(01): 2684
Lewis, David. (1976) ‘Probabilities of Conditionals and Conditional Probabilities’, The Philosophical Review, 85: 297-315.
Lewis (1979) ‘Counterfactual Dependence and Time’s Arrow’. Noûs, 13(4). Reprinted with additional notes in his *Philosophical Papers vol II* (OUP, 1986).
Lewis, David. (1981a) ‘Why Ain’cha Rich?’, Noûs, 15(3): 377-80.
Lewis, David. (1981b) ‘Causal Decision Theory’, Australasian Journal of Philosophy, 59: 530.
McGee, Vann. (1989) ‘Conditional Probabilities and Compounds of Conditionals’, The Philosophical Review, 98(4): 485-541.
Moss, Sarah. (2013). ‘Subjunctive Credences and Semantic Humility’ Philosophy and Phenomenological Research vol. 87, no. 2 (2013): 251-78.
Nozick, Robert. (1969) ‘Newcomb’s Problem and Two Principles of Choice’, in Nicholas Rescher, ed., Essays in Honor of Carl G. Hempel (Dordrecht: Reidel).
Schwarz (unpublished). ‘Subjunctive conditional probability’.
Skyrms, Brian. (1994) ‘Adams conditionals’ in Ellery Eels and Brian Skyrms, eds, Probability and Conditionals - Belief Revision and Rational Decision (Cambridge: Cambridge University Press).
Stalnaker, Robert. (1999) Context and Content - Essays on Intentionality in Speech and Thought (Oxford: Oxford University Press).
van Fraassen, Bas. (1976) ‘Probabilities of Conditionals’, in W. Harper and C. Hooker, eds, Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science (Dordrecht: Reidel).

Williams, J.R.G. (2012) Counterfactual triviality: A Lewis-impossibility proof for counterfactuals Philosophy and Phenomenological Research 85(3).
Williams, J.R.G. (2008) Chances, counterfactuals and similarity; Philosophy and Phenomenological Research, vol 77(2).
Williamson, Timothy. (2007) 'Knowledge of Metaphysical Modality', in The Philosophy of Philosophy (Oxford: Blackwell).