Counterfactual Decision Theory Is Causal Decision Theory

J. DMITRI GALLOW

There is confusion about the sense in which causal decision theory is *causal*. In its earliest formulations, the theory utilised counterfactuals, with scant mention of causation. Subsequent authors introduced the terminology 'causal decision theory', and foregrounded causation, but there remains exegetical disagreement about the role of causation and counterfactuals in these authors' theories. The issue is so vexed that entire articles have been dedicated to the question: 'what is the "cause" in causal decision theory?'.[1]

Most recently, Hedden (2023) argues that 'counterfactual decision theory is importantly different from, and superior to, causal decision theory, properly so-called'. My thesis is that Hedden is mistaken. A careful reading of the sacred texts reveals that counterfactual decision theory is not a competitor to, but rather a version of, causal decision theory—the most popular version by far. I will argue that all of the founding fathers of causal decision theory—Stalnaker (1981), Gibbard & Harper (1978), Lewis (1981), Skyrms (1980, 1982, 1984), Sobel (1994), and Joyce (1999)—endorse counterfactual decision theories. While there are important differences between these authors, the differences concern how they interpret those counterfactuals, and how much they take to be counterfactually determinate, and not whether they rely on counterfactuals.

In §1, I will review the causal decision theories of Stalnaker, Gibbard & Harper, Lewis, Skyrms, Sobel, and Joyce. There, I will present textual evidence that all of these theories are counterfactual. In §2, I will turn to the exegetical question of why these theories were called 'causal'. In §3, I will address Hedden's objections, and show that none of them threaten causal decision theory, properly understood.

1. Causal Decision Theories

There have been a variety of subtly different decision theories defended under the umbrella of 'causal decision theory'. With the exception of Skyrms, they have all been explicitly counterfactual—and even Skyrms accepts a counterfactual reformulation of his theory.

Final Draft; forthcoming in *Pacific Philosophical Quarterly*.
Word Count: 11,019 in body; 1027 in appendix

1. Hitchcock (2013)

In broad outline, the differences between causal decision theories concern how much they take to be counterfactually determinate. The earliest and simplest theories—from Stalnaker and Gibbard & Harper—assume that is always determinate which outcome would result, were you to choose any available option. Lewis and Skyrms assume less. They assume only that it is determinate what the *chance* of any outcome would be, were you to choose any option. Sobel and Joyce assume even less. They allow that even the counterfactual chances could be indeterminate.

1.1. *Counterfactual Determinacy*

Stalnaker (1981) and Gibbard & Harper (1978) advise you to select an act from the set of available acts, $\mathscr{A}$, which maximizes U_1,

$$U_1(A) \stackrel{\text{def}}{=} \sum_{O \in \mathscr{O}} P(A \mathrel{\Box\!\!\rightarrow} O) \cdot V(O) \tag{1}$$

Here, 'P' is your subjective probability function, '$\Box\!\!\rightarrow$' is the counterfactual conditional, $\mathscr{O}$ is the set of potential outcomes, and $V(O)$ is the degree to which you value or desire the outcome O.[2]

This theory works so long as we assume

Counterfactual Determinacy For every available act $A \in \mathscr{A}$ and every world $W \in \mathscr{W}$, there is an outcome $O_A \in \mathscr{O}$ such that the counterfactual $A \mathrel{\Box\!\!\rightarrow} O_A$ is true at W.[3]

To see how things go wrong if counterfactual determinacy is false, suppose I will flip a coin, and I offer you a \$1 bet on whether the coin lands heads for the price of \$100. Suppose that both the counterfactuals 'you take the bet $\Box\!\!\rightarrow$ you lose \$99' and 'you take the bet $\Box\!\!\rightarrow$ you lose \$100' are false at every possible world, since it's indeterminate whether you'd win or lose the bet, were you to take it. Then, your probability in both of those counterfactuals will be zero. So the U_1-value of taking the bet will be zero, too. So the theory would tell you, incorrectly, that it is permissible to take the bet.

2. Instead of a partition of *outcomes*, Stalnaker (1981) uses a partition of *states*. If we assume that outcomes are act-state conjunctions, this is equivalent to (1), since $A \mathrel{\Box\!\!\rightarrow} S$ is the same proposition as $A \mathrel{\Box\!\!\rightarrow} AS$.
3. Counterfactual determinacy is closely related to, but slightly different from, counterfactual excluded middle (CEM). Consider a simple decision with two possible outcomes: O and $\neg O$. Then, CEM says that the disjunction $(A \mathrel{\Box\!\!\rightarrow} O) \vee (A \mathrel{\Box\!\!\rightarrow} \neg O)$ is necessarily true. Whereas counterfactual determinacy says that, necessarily, either $A \mathrel{\Box\!\!\rightarrow} O$ is true or $A \mathrel{\Box\!\!\rightarrow} \neg O$ is true. Assuming a classical semantics, CEM and counterfactual determinacy will be equivalent. But with a supervaluationist semantics, you could say that the *disjunction* $(A \mathrel{\Box\!\!\rightarrow} O) \vee (A \mathrel{\Box\!\!\rightarrow} \neg O)$ is necessarily true, even though neither disjunct is even possibly true (see Stalnaker, 1980). If we favor a supervaluationist semantics, Stalnaker, Gibbard, and Harper's theory will need more than CEM; it will need the stronger assumption of counterfactual determinacy.

1.2. *Counterfactual Chance Determinacy*

The simplest theory works if counterfactual determinacy is true, but Lewis rejected counterfactual determinacy.[4] So he defended a more general version of causal decision theory. Lewis (1981) says to maximize the function U_2,

$$U_2(A) \stackrel{\text{def}}{=} \sum_{K \in \mathcal{K}} P(K) \cdot V(AK) \tag{2}$$

Here, '$\mathcal{K}$' is the set of *causal dependency hypotheses*, and Lewis's function V is defined over arbitrary propositions. Lewis starts by defining V on a set of possible worlds, $\mathcal{W}$. V is then lifted to arbitrary propositions by stipulating that, for any proposition X, $V(X) \stackrel{\text{def}}{=} \sum_{W \in \mathcal{W}} P(W \mid X) \cdot V(W)$. It then follows that, for any proposition X and any partition $\mathcal{Z}$, $V(X) = \sum_{Z \in \mathcal{Z}} P(Z \mid X) \cdot V(XZ)$. Skyrms also endorses U_2-maximisation, though his characterisation of the dependency hypotheses is different from Lewis's (more on this below).

Lewis's formulation (2) appears different from (1), but given a few assumptions, they turn out to be equivalent. Assume counterfactual determinacy and[5]

Value-level Outcomes Each outcome settles everything you care about, so that, for each outcome $O \in \mathcal{O}$, and any two worlds $W, W^* \in O$, $V(W) = V(W^*)$.

Outcomes as Act-State Conjunctions There is a partition of *states*, $\mathcal{S}$, such that each outcome $O \in \mathcal{O}$ is equivalent to a conjunction AS, for some available act $A \in \mathcal{A}$ and some state $S \in \mathcal{S}$.

With these assumptions, we can let each dependency hypothesis $K \in \mathcal{K}$ be a conjunction of counterfactuals specifying which outcome would obtain, were you to choose each act, $K = \bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} O_{A,K}$ (where $\mathcal{A}$ is the set of available acts and $O_{A,K}$ is the unique outcome which would result, were you to choose A, according to K). With this stipulation, our assumptions imply that $U_1(A) = U_2(A)$, for each $A \in \mathcal{A}$. (This is proven in Lewis's 1981 article, but I provide a proof in the appendix to spare the reader a trek through the literature.)

Lewis rejected counterfactual determinacy, but he accepted a probabilistic analogue. While he denied that outcomes were counterfactually determinate, he did think that the *chances* of various outcomes were counterfactually determinate. Some notation and terminology: I'll use '$\mathcal{C}h_{\mathcal{Z}}$' for the definite description 'the objective chance distribution over the cells of the partition $\mathcal{Z}$ at the moment of choice'. If the partition is the set of all possible worlds, then I'll just write '$\mathcal{C}h$'. (That is, $\mathcal{C}h = \mathcal{C}h_{\mathcal{W}}$.) Lewis called a partition of states, $\mathcal{S}$, *rich* iff, for each $S \in \mathcal{S}$, each act $A \in \mathcal{A}$, and each dependency hypothesis $K \in \mathcal{K}$, $V(ASK) = V(AS)$. That is, for a rich partition of

4. See Lewis (1973*b*)
5. For justification of these assumptions, see the discussion in Joyce (1999, §2.2).

states, 'The *AS*'s describe outcomes of options so fully that the addition of a dependency hypothesis tells us no more about the features of the outcome that matter to the agent'.[6] Then, Lewis assumed that

Counterfactual Chance Determinacy There is some rich partition of states $\mathcal{S}$ which describes occurrences mereologically distinct from your choice, and such that for every act $A \in \mathcal{A}$ and every world $W \in \mathcal{W}$, there is a chance distribution ch_A which makes the counterfactual $A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch_A$ true at W.

Assuming that we identify the propositions $A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}}(S) = 1$ and $A \mathrel{\Box\!\!\rightarrow} S$, this assumption is strictly weaker than the assumptions we made before (counterfactual determinacy, value-level outcomes, and outcomes as act-state conjunctions).

Counterfactual chance determinacy allowed Lewis to take each dependency hypothesis to be a conjunction of counterfactuals specifying what the objective chances would be, were you to choose any of the available acts. That is, for Lewis, a causal dependency hypothesis takes the form $K = \bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch_A$, where $\mathcal{S}$ is a rich partition specifying occurrences distinct from your act, and each 'ch_A' is an objective chance distribution over $\mathcal{S}$—the objective chance distribution over $\mathcal{S}$ that K says would obtain, were you to choose A. Lewis called a set of counterfactuals of the form $A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch_A$, with one such counterfactual for each available act, a 'probabilistic full pattern'. He claimed that the conjunction of counterfactuals in any such set is a dependency hypothesis:

> I claim that the conjunction of the counterfactuals in any probabilistic full pattern is a causal dependency hypothesis.[7]

So any probabilistic full pattern gives us a dependency hypothesis. Can we get *every* dependency hypothesis in this way? Lewis expresses some reservations about this stronger thesis,[8] but despite these misgivings, he signs on for the stronger view that the dependency hypotheses are precisely the conjunctions of probabilistic full patterns, and that 'we have succeeded in capturing all the dependency hypotheses by means of counterfactuals'.[9] He summarizes his article as advancing two theses:

> My main thesis is that we should maximise expected utility calculated by means of dependency hypotheses [*i.e.*, U_2]...My subsidiary thesis...is that

6. Lewis (1981, p. 15)
7. Lewis (1981, p. 26)
8. He doesn't clearly explain the reason for these doubts, but I think the best reading is that he was concerned that even the chances might fail to be counterfactually determinate. Immediately after saying 'I shall assume [that we have succeeded in capturing all the dependency hypotheses by means of counterfactuals], not without misgivings', he says: 'That means accepting a special case of Conditional Excluded Middle'. (Lewis, 1981, p. 27) It's most natural to read the second sentence as an explanation of the misgivings mentioned in the first.
9. Lewis (1981, p. 27)

the dependency hypotheses are exactly the conjunctions of probabilistic full patterns.[10]

Given Lewis's subsidiary thesis, U_2 reduces to U_1 in many situations. In particular, whenever

Counterfactual Determinism For each act $A \in \mathscr{A}$ and each world $W \in \mathscr{W}$, there is a state $S_A \in \mathcal{S}$ such that $A \mathrel{\Box\!\!\rightarrow} \mathscr{C}h_{\mathcal{S}}(S_A) = 1$ is true at W.

Again, equating $A \mathrel{\Box\!\!\rightarrow} S$ and $A \mathrel{\Box\!\!\rightarrow} \mathscr{C}h(S) = 1$, and assuming Lewis's thesis about causal dependency hypotheses, we will have $U_1(A) = U_2(A)$ for each $A \in \mathscr{A}$ whenever counterfactual determinism holds. (See the appendix for a proof.)

Skyrms (1980, 1982, 1984) joins Lewis in advocating U_2-maximisation, but he did not want to characterize dependency hypotheses *directly* in terms of counterfactuals, since he wanted to 'avoid unanalyzed counterfactual machinery as much as possible',[11] due to 'the ambiguity of subjunctive conditionals'.[12] He instead takes a dependency hypothesis to be a maximally specific proposition about 'the factors outside our influence at the time of decision which are causally relevant to the outcome of our actions'.[13] To distinguish Lewisian dependency hypotheses from Skyrmsian dependency hypotheses, I'll use 'K' for Lewisian conjunctions of probabilistic full patterns, and 'k' for a Skyrmsian dependency hypothesis about relevant factors beyond your control.

Skyrms's k's are very nearly Lewis's K's, if we assume that what it is for a factor to be outside your influence is for its chance to be counterfactually independent of your choice (I'll have more to say on the analysis of 'causal influence' in the next section). But even if we think that one of their dependency hypotheses holds fixed more than the other's, it won't make any difference to the calculation of U_2. For Skyrms assumed that one of his dependency hypotheses, k, in conjunction with your choice, A, would determine a chance distribution over outcomes, $ch_{A,k}$. I'll suppose that the relevant kind of determination is metaphysical necessitation, so that what it is for k and A to determine the chance function $ch_{A,k}$ is for $\mathscr{C}h = ch_{A,k}$ to be true at every world in Ak. Because he accepted counterfactual chance determinacy, (and *modus ponens*), Lewis also held that one of *his* determinacy hypotheses, K, in conjunction with your choice, A, would metaphysically necessitate a particular chance distribution, $ch_{A,K}$. So both Skyrms's $\mathscr{k}$ and Lewis's $\mathscr{K}$ are what we can call 'chance determining' partitions.

Suppose that you have no inadmissible evidence and that you satisfy Lewis's *principal principle*, so that your probability for any proposition X, given any admissible proposition E and given that the chances at the moment of choice are ch, is $ch(X)$. Then, we can say that a chance determining partition $\mathscr{K}$ is *admissible* iff every cell of

10. Lewis (1981, p. 27)
11. Skyrms (1980, p. 139, fn 16)
12. Skyrms (1980, p. 132)
13. Skyrms (1980, p. 133)

it is admissible. That is, $\mathcal{K}$ is admissible iff, for every $A \in \mathcal{A}$, every $K \in \mathcal{K}$, and every proposition $X \subseteq \mathcal{W}$, $P(X \mid AK) = ch_{A,K}(X)$.

In the appendix, I show that, if you have two chance-determining admissible partitions, $\mathcal{K}$ and $\mathscr{k}$, such that one is a refinement of the other (meaning that every cell of the one entails some cell of the other), then the value of U_2 is unaffected by which of $\mathcal{K}$ and $\mathscr{k}$ you use to calculate it. That is, even if one of their partitions holds more fixed than the other's, so long as both Lewis's $\mathcal{K}$ and Skyrms's $\mathscr{k}$ are admissible chance-determining partitions, we will have

$$\sum_{K \in \mathcal{K}} P(K) \cdot V(AK) = \sum_{k \in \mathscr{k}} P(k) \cdot V(Ak)$$

So Lewis's and Skyrms's theories will agree.

Skyrms came to the same conclusion via a different route. He didn't want to use *unanalysed* counterfactuals, but he offered an analysis of counterfactuals in terms of objective chance. He said that the probability of a counterfactual 'if I were to choose A, it would be that C' is equal to your subjective expectation of the chance of C, conditional on you choosing A, and conditional on the factors beyond your influence at the moment of choice. (This is sometimes called 'Skyrms's Thesis'.) And he likewise says that a counterfactual with a chance consequent 'if I were to choose A, the chance of C would be x' is true iff A, together with the factors beyond your influence, determine a chance of x for C. Given this analysis, Skyrms's theory will agree with Stalnaker's, Gibbard & Harper's, and Lewis's.

So, even Skyrms was happy to accept a counterfactual formulation of his theory. He says:

> I can...accept both Gibbard-Harper decision theory and Lewis's generalization thereof to conditionals with chancy consequents...given my analysis of the subjunctives involved[14]

The disagreement between Skyrms and Lewis wasn't whether to use counterfactuals, but rather which kinds of counterfactuals to use. Lewis's counterfactuals were English counterfactuals on their standard reading, whereas Skyrms's counterfactuals were disciplined regimentations of ordinary English counterfactuals—explicitly stipulated to hold fixed factors other than your choice and the things causally influenced by your choice, even if English counterfactuals do not always work this way.[15] Theirs was a disagreement over what to take as primitive and what to take as defined, not over whether counterfactuals were relevant to rational choice.

14. Skyrms (1980, p. 139)
15. The same is true of interventionist versions of causal decision theory (see, for instance, Meek & Glymour 1994, Hitchcock 2016, Stern 2017, 2018, and Gallow forthcoming*b*, appendix A.1). They use interventionist counterfactuals, which are explicitly stipulated to hold fixed variables besides your choice and variables causally influenced by your choice.

1.3. *Counterfactual Chance Indeterminacy and Primitive Imaging*

Sobel (1986, 1994) and Joyce (1999) formulate causal decision theory with the aid of an *imaging function*. An imaging function takes as input a possible world $W \in \mathcal{W}$ and an act $A \in \mathcal{A}$ and returns a probability distribution, W^A, such that $W^A(A) = 1$. Using this imaging function, they define your probabilities *imaged* on A,

$$P^A(C) \stackrel{\text{def}}{=} \sum_{W \in \mathcal{W}} W^A(C) \cdot P(W)$$

which Joyce calls a 'causal probability function'. And they advise you to select an available act which maximizes

$$U_3(A) \stackrel{\text{def}}{=} \sum_{W \in \mathcal{W}} P^A(W) \cdot V(W) \tag{3}$$

Both Sobel and Joyce interpret the imaging function counterfactually. On Sobel's interpretation, if $W^A(C) = 1$, then $A \mathrel{\Box\!\!\rightarrow} C$ is true at W. If $W^A(C) = 0$, then $A \mathrel{\Box\!\!\rightarrow} \neg C$ is true at W. If $W^A(C) = x$, for some x between 0 and 1, then the quantified 'might' counterfactual 'If you were to choose A, then it might with a chance of x be that C' is true at W. This interpretation is reflected in the notation Sobel used for the imaging function. Rather than writing '$W^A(C) = x$', Sobel wrote '$A \Diamond_x\!\!\rightarrow C$'. Here is what Sobel says about how this function is to be understood:

> Let '$A \Diamond_x\!\!\rightarrow C$'...say that if it were the case that A, then it *might* be the case that C—might with a chance of x...[16]

Sobel then goes on to define $P^A(C)$ in terms of an expectation of the value of x in $A \Diamond_x\!\!\rightarrow C$, which is just the definition of P^A we gave above.

Joyce also interprets the causal probability function P^A counterfactually:

> I'd like to register my preference for a subjunctive reading of [P^A]...I regard decision making as an essentially subjunctive endeavor; to evaluate an act, in the causally relevant sense, is to ask what would happen were it to be done. Given such an outlook, each causal probability [$P^A(W)$] that appears in [U_3] should capture the decision maker's views about *how likely W would be were A to be performed*...[17]

The formulation (3) looks different from (2), but there is a way of interpreting the imaging function on which they turn out to be equivalent in many circumstances. Some notation: write '$\mathcal{C}h^A_{\mathcal{X}}$' for the chance distribution over $\mathcal{X}$ that would obtain, were you to choose A. And use '$\mathcal{C}h^A_{\mathcal{X},W}$' for the chance function picked out by this

16. Sobel (1986, p. 418, with minor notational changes)
17. Joyce (1999, p. 172).

definite description at the world W. Then, assume counterfactual chance determinacy and both of the following

Images as Counterfactual Chances For each world $W \in \mathscr{W}$ and each act $A \in \mathscr{A}$, $W^A = \mathscr{C}h_W^A$

No Inadmissible Evidence You don't have any inadmissible evidence about the future, you satisfy Lewis's *principal principle*, and the $\mathscr{K}$ partition is admissible.

Then, Sobel and Joyce's theory will be equivalent to Lewis's and Skyrms's, in the sense that $U_3(A) = U_2(A)$, for each $A \in \mathscr{A}$. (Again, this is proven elsewhere, but I provide a proof in the appendix for the reader's convenience.)

Both Sobel and Joyce have reservations about interpreting images as counterfactual chances, but both think the assumption holds in many cases. Sobel writes:

> Difference between chance-conditionals and conditionals with chance-consequents, and between Lewis' theory and mine, are not…great.[18]

Joyce (1999, p. 176) thinks that images are *almost* counterfactual chances, but he harbors concerns about counterfactual chance determinacy. This leads him to say only that $P^A(C)$ should be bounded from below by your subjective expectation of the counterfactual chance C would have, were you to choose A, and bounded from above by one minus your expectation of the counterfactual chance $\neg C$ would have, were you to choose A.[19] Nonetheless, he agrees that images are counterfactual chances 'when determinacy reigns'.[20]

So U_3 is a generalisation of U_2, just as U_2 was a generalisation of U_1. As we've already seen, U_2 reduces to U_1 if we make the assumption of counterfactual determinism. So, if we assume counterfactual determinism while taking images to be counterfactual chances—and assuming that outcomes are value-level act-state conjunctions—then all of these theories will collapse onto each other.

1.4. Hedden's Counterfactual Decision Theory

As I mentioned in the introduction, Hedden (2023) defends counterfactual decision theory over causal decision theory. Let me spend a bit of time discussing the relation-

18. Sobel (1986, p. 419)
19. That is, he requires that

$$\sum_x x \cdot P(\mathscr{C}h^A(C) = x) \leqslant P^A(C) \leqslant 1 - \sum_x x \cdot P(\mathscr{C}h^A(\neg C) = x)$$

Here's a way of guaranteeing that P^A lies in this range: use 'W_A' for the strongest proposition which would be true, were you to choose A at W. Then we could set $W^A(C) = \sum_{W^* \in \mathscr{W}} P(W^* \mid W_A) \cdot \mathscr{C}h_{W^*}(C)$. Assuming that $\mathscr{C}h_W^A(C) = x$ is true iff W_A entails that $\mathscr{C}h(C) = x$, this definition of W^A will satisfy Joyce's constraint.

20. Joyce (1999, p. 176)

ship between what Hedden calls 'counterfactual decision theory' and the theories we have just reviewed.

Hedden's preferred form of counterfactual decision theory is just the theory of Stalnaker, Gibbard, and Harper. This theory says to maximize U_1. So he has no disagreement with them. What are the disagreements between Hedden and the other causalists? Not counterfactual determinacy. While he is personally attracted to counterfactual determinacy (that's why he favors U_1), in §6 of his article, Hedden discusses how counterfactual decision theory could be generalized to accomodate failures of counterfactual determinacy. He makes two suggestions. First, he suggests that we could

> employ a $\mathcal{K}$-partition along the lines of Lewis (1981) and define counterfactual expected utility as [U_2]. But whereas Lewis defines the K's, which he called 'dependency hypotheses', in terms of causation...a counterfactual decision theorist could employ dependency hypotheses defined in terms of counterfactuals, for example, as maximally specific ways that things might be in ways counterfactually independent of your choice.[21]

Lewis's dependency hypotheses *are* defined in terms of counterfactuals, albeit counterfactuals of a certain sort: those whose consequents give chance distributions over the states in $\mathcal{S}$. Hedden's suggestion is to use counterfactuals in a superficially different way. If I understand, the proposal is this: for each world $W \in \mathcal{W}$, let H_W be the strongest proposition which would be true, no matter how you choose. That is, it is the strongest proposition such that the conjunction $\bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} H_W$ is true at W. Then, we could let Hedden's partition of dependency hypotheses be the set $\mathcal{H} = \{H_W \mid W \in \mathcal{W}\}$. By weak centering, $W \in H_W$. It's a substantial assumption that $\mathcal{H}$ is a partition, but it is an assumption Hedden appears willing to make. Then, it must be that, whenever $W^* \in H_W$, $H_{W^*} = H_W$. Else, H_W and H_{W^*} would overlap at W^*, and $\mathcal{H}$ would not form a partition. But this means that each of the counterfactuals $A \mathrel{\Box\!\!\rightarrow} H_W$ are counterfactually independent of your choice. So their conjunction is also counterfactually independent of your choice.[22] So H_W entails the conjunction $\bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} H_W$. Since $\mathcal{A}$ is a partition, the conjunction entails H_W.[23] So each of Hedden's dependency hypotheses H is exactly the conjunction $\bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} H$.

The same is true of Lewis's dependency hypotheses, the K's. Let K_W be the Lewisian dependency hypothesis which is true at W. Since Lewis's dependency hypotheses are act-independent,[24] the conjunction $\bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} K_W$ is true at W. By weak centering, $W \in K_W$. Assuming they form a partition, $W^* \in K_W$ implies

21. Hedden (2023, p. 755–756)
22. Here, I assume that $A \mathrel{\Box\!\!\rightarrow} C$ and $A \mathrel{\Box\!\!\rightarrow} D$ imply $A \mathrel{\Box\!\!\rightarrow} CD$.
23. This follows from *modus ponens* and proof by cases.
24. 'Dependency hypotheses are "act-independent states" in a causal sense, though not necessarily in the probabilistic sense' (Lewis, 1981, p. 13). More on this below.

$K_{W^*} = K_W$. So each counterfactual $A \mathrel{\Box\!\!\rightarrow} K_W$ is counterfactually independent of your act. So their conjunction is counterfactually independent of your act. So K_W entails $\bigwedge_{A \in \mathscr{A}} A \mathrel{\Box\!\!\rightarrow} K_W$. And since $\mathscr{A}$ is a partition, the conjunction entails K_W. So each Lewisian dependency hypothesis K is exactly the conjunction $\bigwedge_{A \in \mathscr{A}} A \mathrel{\Box\!\!\rightarrow} K$.

Unlike Hedden's H's, Lewis's K's need not be the *strongest* thing that is counterfactually independent of your act. They only say things about how the chances of the states in the rich partition $\mathscr{S}$ do and do not causally depend upon how you choose. Nonetheless, even if $\mathscr{H}$ is not the same partition as $\mathscr{K}$, it will be a refinement of it (in the sense that every $H \in \mathscr{H}$ will entail some $K \in \mathscr{K}$.) For if K_W is counterfactually independent of your choice at W, then it must be entailed by the *strongest* thing counterfactually independent of your choice at W. Assuming counterfactual chance determinacy, so that $\mathscr{K}$ *is* a partition, and assuming that $\mathscr{H}$ is an *admissible* partition, there will be no difference between this theory and Lewis's. For, as I mentioned above, and as I prove carefully in the appendix, the value of U_2 is the same no matter which of two partitions we use to calculate it, so long as both are admissible chance-determining partitions and one refines the other.

However, conditional on rejecting counterfactual determinacy, Hedden denies counterfactual chance determinacy, too. He says that, 'in so far as we are inclined to doubt [counterfactual determinacy], we'll also think that sometimes there are no precise chance facts that would obtain were you to perform a given action'.[25]

His other suggestion for generalizing U_1 to handle failures of counterfactual determinacy is to use a counterfactual imaging function to specify your probabilities *imaged* on the performance of A, P^A, and to maximize $U_4(A) = \sum_{O \in \mathscr{O}} P^A(O) \cdot V(O)$. U_4 differs from U_3 only in that the summation is being taken over outcomes instead of worlds. If we assume outcomes are value-level, then $U_4(A)$ will be equal to $U_3(A)$, for every $A \in \mathscr{A}$ (I offer a proof of this fact in the appendix). So this suggestion is to adopt the theory of Joyce and Sobel.

At one point, Hedden includes Joyce's theory among the 'standard formalizations of causal decision theory',[26] but later, he says that Joyce's theory is in fact a version of counterfactual decision theory, and that 'my only objection is terminological...he should not self-identify as a "causal" decision theorist.'[27] So I'll assume that Hedden has no substantive disagreement with either Joyce or Sobel. On his use of the term 'causal decision theory', Joyce and Sobel do not count as causal decision theorists.

In sum, there appears to be no substantive disagreements between Hedden and Stalnaker, Gibbard & Harper, Sobel, and Joyce. I can identify two disagreements between Hedden and Lewis & Skyrms: firstly, unlike Lewis, Hedden does not ex-

25. Hedden (2023, p. 755). Hedden is talking about the principle of counterfactual excluded middle, which implies counterfactual determinacy together with bivalence.
26. Hedden (2023, p. 737)
27. Hedden (2023, p. 741)

plicitly take outcomes to be act-state conjunctions, and does not require that states be mereologically distinct from how you choose. Secondly, unlike Lewis & Skyrms, Hedden is unwilling to tolerate a middle-ground position on counterfactual determinacy. According to him, if there's no determinate fact-of-the-matter about which outcome would obtain, were you to choose A, then there's likewise no determinate fact-of-the-matter about what the *chance* of an outcome obtaining would be, were you to choose A. Hedden argues that this first difference makes a difference in cases where you care about which of your acts you select. In §3.2 below, I'll argue that this is incorrect. The second difference is a genuine one. It's nonetheless worth noting that Hedden's concerns are shared by many causalists. Lewis himself expressed misgivings about the assumption of counterfactual chance determinacy. While he endorsed his 'subsidiary thesis' that all dependency hypotheses are conjunctions of probabilistic full patterns, he said that he did so 'not without misgivings',[28] and that he put it forward 'much more tentatively' than his main thesis, that we should be maximizing U_2, given some characterisation of causal dependency hypotheses. He writes that 'If ever we must retract our assumption that there is a probabilistic full pattern for each world [i.e., counterfactual chance determinacy], the two approaches [his and Sobel's] will separate and we may need to choose; but let us cross that bridge if we come to it.'[29] Joyce (1999) argued that we should cross that bridge, and that we should favour Sobel's approach. Here is his example:

> Suppose that you were to have one more child than you actually ever do. What are the objective chances that she would grow up to be a dentist? Answer: There is no answer because the actual facts do not make any chance assignment reasonable.[30]

So there's some disagreement between Hedden and Lewis/Skyrms. But it's not a disagreement about the role that counterfactuals have to play in rational deliberation. Nor is it a disagreement about whether the theory of Lewis/Skyrms is correct. For Hedden accepts counterfactual determinacy, and counterfactual determinacy implies that $U_1 = U_2$—as long as outcomes are act-state conjunctions and value-level. Given these assumptions, Hedden should think that the theory of Lewis/Skyrms is correct. Rather, the substantive disagreement is about whether we should accept the theory of Lewis/Skyrms *if* we reject counterfactual determinacy, or whether we should instead opt for the theory of Sobel/Joyce.

In the next section, I will turn to the question of why causal decision theories were so-called. In §3, I'll consider Hedden's objections to causal decision theory. There, I will argue that, properly understood, none of the theories from §1 are threatened by these objections.

28. Lewis (1981, p. 27)
29. Lewis (1981, p. 28)
30. Joyce (1999, p. 172)

2. Why Are Causal Decision Theories So-Called?

The founding fathers of causal decision theory—Stalnaker, Gibbard & Harper, Skyrms, Lewis, Sobel, and Joyce—all accepted counterfactual formulations of their theories. Given this, why isn't the theory called 'counterfactual decision theory'? That could easily have been the name they settled on—and, after all, a rose by any other name would smell as sweet. But let me say something about why some of the founding fathers favored this name, and in what sense the theory is causal.

2.1. Causal Influence and Causal Control

Many have thought that counterfactuals admit of multiple readings. For instance, consider a perfect predictor version of Newcomb's problem[31], and imagine that you two-boxed, walking away with $1000. In this context, many of us can hear a true reading of both (4) and (5).[32]

4) If you had one-boxed, you would have been predicted to one-box, so you would have gotten $1,000,000.

5) If you had one-boxed, you would (still) have been predicted to two-box, so you wouldn't have gotten anything.

Causalists want their theory to tell you to two-box in Newcomb's problem, so it's important that they interpret their counterfactuals along the lines of (5). Counterfactuals like (4) say something about the necessary causal precursors of the antecedent; whereas counterfactuals like (5) hold fixed everything except your choice and things causally influenced by your choice. For this reason, claims like (5) are often called 'causal counterfactuals'. Causal counterfactuals reveal which parts of the world are under your causal control and which are not. In causal decision theories, it is important that the counterfactuals being used are *causal* counterfactuals.

Hedden doesn't think that counterfactuals like (5) deserve the name 'causal'. He writes:

> why should we accept that standard counterfactuals [i.e., counterfactuals like (5)] are causal? Standard counterfactuals and causation certainly have something to do with each other...Perhaps even some counterfactual analysis of causation is correct. But even defenders of such analyses, like Lewis (1973*a*), deny that the relationship is anywhere near as simple as the equivalence of '*A* causes *B*' and 'if *A* were the case, *B* would be the case'.[33]

31. See Nozick (1969)
32. Cf. Horgan (1981).
33. Hedden (2023, p. 733)

Hedden is correct that the relationship between counterfactuals like (5) and token causation is vexed and complicated. But that relationship isn't the one which warrants the name. Whatever their relationship to *token causation*, counterfactuals like (5) are intimately and directly related to another important kind of causal relation which I'll call *causal influence*. When we talk about token causation, we usually use the verb 'to cause' and we relate token happenings like 'the battle' and 'the window breaking', or noun phrases with infinitival clauses, like 'the window to break' or 'the cotillion to abruptly end'. We say things like 'The window breaking caused the battle' and 'the battle caused the cotillion to abruptly end'. But these aren't the only kinds of causal claims we make. We also make claims about causal influence. When we talk about influence, we usually use the verbs 'to affect' or 'to influence', and relate interrogative clauses like 'whether we catch the bus', 'how you choose' and 'when we arrive'. We say things like 'whether we catch the bus influences when we arrive' and 'how you choose affects what you win'.

These two different causal relations have two different roles to play. Token causation plays a role in moral responsibility. You can only be appropriately blamed and praised for your behaviour and its effects. Causal influence, on the other hand, has a role in play in our thinking about what is and is not under your control. You have control over whether ϕ iff how you choose influences whether ϕ is true or false. And some outcome is under your control iff you have control over whether it happens. In a slogan: while you're responsible for your effects, you have control over what you affect.

We can illustrate the difference between these two relations with one of the decisions from Hedden (2023):

Overdetermination You want the window broken. You don't care how it happens; you just want it broken. You see that I've just thrown my rock at the window. If you do nothing, my rock will break the window. If you throw, your rock will hit mine and deflect it, and then your rock will hit the window, causing it to break. But it will also cause some unwanted energy expenditure.[34]

In this decision, whether you throw does not affect whether the window breaks or not. So you have no control over whether it breaks. The window's breaking is not under your control.[35] However, if you throw, then your throw will cause the window to break—that is to say, the breaking will be an effect of the throw. If you throw, you will be morally responsible for the window breaking, it will be fitting to blame you for it breaking, and you will have a duty to repair the window.

34. Hedden (2023, p. 732)

35. In a realistic version of this case, you could throw the rock at a different angle so that the window didn't break. In that case, you would have some control over whether the window breaks. For the purposes of the example, let's imagine that you have only two options, throwing and not throwing, and that the window will certainly shatter no matter what you do.

(It was only recently that philosophers started carefully distinguishing these two kinds of causal relations, so the terminology has not been standardized. Several just use 'causation' or 'causality' for what I'm calling 'causal influence'. What I'm calling 'token causation' is elsewhere called 'singular causation', 'event causation', and 'actual causation'. Often, authors are only focused on one or the other of these relations, and reserve 'causation' for it—this has no doubt led to some confusion. The distinction seems to have come into focus with the advent of causal modelling techniques, where it becomes necessary to distinguish the causal relations holding between variables—which correspond to the interrogative clauses related by causal influence—and the actual values of those variables—which correspond to the relata of token causal claims.)

Lewis distinguished between a relation that he called *causal dependence* and the relation of token causation. Causal dependence is (causal) counterfactual dependence of the chance of one event on another, mereologically distinct, event. That is, for any two distinct events, *c* and *e*, *e* causally depends upon *c* iff *e* would have a higher chance of happening were *c* to occur than it would have were *c* to not occur.[36] This relation of causal dependence is very nearly the relation of causal influence. Their relata are different; causal influence relates interrogative clauses, and causal dependence relates events. But when we are interested in whether an event occurs, there is a straightforward connection. Whether *c* happens causally influences whether *e* does iff *e* causally depends upon *c*. And the relation of causal dependence is what Lewis used in his version of causal decision theory. He took a dependency hypothesis to be a hypothesis about how states, mereologically distinct from your choice, do and do not causally depend upon how you choose. Thereby, a causal dependency hypothesis reveals what you do and do not have causal control over.

You might think that causal influence and causal dependence come apart in some cases. For instance, if we assume that the laws of nature are deterministic, then many theories of counterfactuals force us to choose between 'had you done otherwise, the initial conditions would have been different' and 'had you done otherwise, the actual laws of nature would have been violated'. It's most natural to think both that you don't influence the initial conditions and that you don't influence the laws of nature. So you might think this shows us that something—either the laws or the initial conditions—counterfactually depends upon how you choose, even though you do not causally influence that thing.[37] In that case, 'causal decision theory' might not be aptly named. Though perhaps it could still warrant the name in virtue of the general—if not exceptionless—connection between counterfactuals and causal influence? (As an aside, I disagree that causal influence and causal dependence come apart in these cases,

36. See Lewis (1986*c*) for the definition of causal dependence and Lewis (1986*a*) for more on the mereology of events.
37. Cf. Dorr (2016)

for two reasons. Firstly, I reject both of the relevant counterfactuals when they are given a causal reading.[38] Secondly: were I persuaded to accept one of them, I would *thereby* be persuaded that your choice causally influences either the initial conditions or the laws. By definition, causal counterfactuals about how you choose hold fixed things which are not causally influenced by how you choose. So if I were persuaded that the laws causally counterfactually vary with your choice, I would thereby be persuaded that your choice causally influences the laws.)

So I claim that the 'causal' in 'causal decision theory' does not refer to the relation of token causation, but rather the relation of causal influence, which is very nearly causal dependence. This isn't a novel exegetical claim. Hitchcock (2013) asks the titular question 'What is the 'cause' in causal decision theory?', and gives the same answer:

> The specific relation that [those] concerned with preemption and overdetermination cases have been trying to analyze is *actual causation*. It has been called by other names: 'singular causation', 'token causation', 'event causation', and just plain 'causation'...the relation that figures in formulations of causal decision theory is not actual causation...Lewis's relation of causal dependence is deserving of the epithet 'causal'...because it is a genuine causal relation in its own right. It is just this notion of causal dependence that is captured...in CDT.[39]

The founding fathers occasionally misspoke and used the language of token causation when describing their views. For instance, Skyrms (1982) opens with 'practical reason should evaluate actions in terms of their *causal consequences*'. It's natural to read 'causal consequences' as 'effects'. Similarly, Lewis (1981) opens with: 'Decision theory in its best-known form manages to steer clear of the thought that what's best to do is what the agent believes will most tend to cause good results.' On its most natural reading, 'tend to cause good results' means 'tend to have good results as its effects'.

Much more often, however, they used the language of causal influence and control. Skyrms (1980) describes his causal dependency hypotheses as 'maximally specific specifications of the factors outside our influence at the time of decision'.[40] Skyrms (1984) opens his chapter on causal decision theory by attributing to Aristotle the view that 'we deliberate about factors which our actions may affect'.[41] At the end of his opening paragraph, Lewis (1981) condemns non-causal decision theory for 'managing the news so as to get good news about matters which you have no control over'.[42] And he says that causal dependency hypotheses 'specify [the decision maker's] influence

38. See Gallow (forthcoming*a*).
39. Hitchcock (2013, p. 140–141)
40. Skyrms (1980, p. 133)
41. Skyrms (1984, p. 63)
42. Lewis (1981, p. 5)

over other things'.[43] When discussing a particular dependency hypotheses which says how Bruce's purring causally depends upon his choices, he writes that it 'says that loud and soft purring are within my influence' and that it 'specifies the extent of my influence, namely full control'.[44]

Other of the founding fathers have similar slip-ups where they talk about token causation rather than causal influence. And over the years, some expositors have followed suit, describing causal decision theory using the language of token causation, rather than causal influence. For the most part, however, they stick to the language of control and causal influence. The occasional slip-ups are understandable, for two reasons. In the first place, at the time causal decision theory was being developed, philosophers had not carefully distinguished causal influence from token causation. In the second place, cases of overdetermination aren't the norm. Normally, all of the effects of your choice are under you control. And normally, you don't care about your choice intrinsically. And in those circumstances, causal decision theorists will say that you should make the choice whose expected effects are best. So it's not terribly misleading to summarize the view with a slogan like this, even though cases of overdetermination show that the slogan is not strictly correct.

In sum: causal decision theories are not so-called because they tell you to only care about the effects of your choices. Rather, they are so-called because they tell you to care only about things you are in a position to causally affect, and to ignore things which are outside of your causal control.

2.2. Accomodating Reasonable Disagreement About Causal Influence

There's another reason Lewis, Skyrms, and Joyce stuck to 'causal decision theory' rather than 'counterfactual decision theory': they wanted to keep their faction from splintering into schisms, and so wanted to accommodate reasonable disagreement about how to analyse causal influence. As Joyce explains: 'causal decision theorists have done a great deal of work aimed at minimizing the differences among their formulations'.[45] Finding common ground was Lewis's primary aim in his 1981 article—in the introductory paragraphs, he assures readers that '[t]he situation is not the chaos of disparate approaches that it might seem' (p. 5). To this end, they sought to formulate their theories in a neutral, inclusive way. All could agree that, when deciding, you should ignore goods you are not in a position to causally influence, even if they disagree about how we should analyse the relation of causal influence. Lewis and Joyce's preferred analysis used causal counterfactuals, but they wanted to allow others to understand causal influence and causal control differently. So they wanted their theories to be able to accommodate those alternative interpretations.

43. Lewis (1981, p. 13)
44. Lewis (1981, p. 21)
45. Joyce (1999, p. 171)

Both Lewis and Skyrms wanted to accommodate any reasonable disagreement about causal influence as disagreement about how to understand causal dependency hypotheses. Skyrms tells us that

> it is in the further analysis of [causal dependency hypotheses] that causal decision theorists part ways...those who have counterfactual machinery use that machinery to give an account of [depedency hypotheses]; reductionists move to implement their reductionist program, etc.[46]

In a similar spirit, when Lewis introduces the notion of a causal dependency hypothesis in §5 of his article, he does not introduce it immediately as a conjunction of the counterfactuals in a probabilistic full pattern. Rather, he characterizes it more neutrally as 'a maximally specific proposition about how the things [the decision maker] cares about do and do not depend causally on his present actions'. And he stresses that, while

> there are certainly differences [between causalists] about the nature of dependency hypotheses...these are small matters compared to our common advocacy of utility maximising as just defined [*i.e.*, U_2].

Joyce wanted to accomodate reasonable disagreement about causal influence as disagreement about the definition of a 'causal probability function', P^A. To this end, he showed how to define P^A in terms of either a counterfactual imaging function or in terms of a partition of causal dependency hypotheses. He notes that you could specify such a partition in terms of the 'causally relevant factors' from a probabilistic theory of causal influence, or in terms of conditional chances, or in terms of counterfactual dependence. As Joyce puts it:

> Those who go in for probabilistic analyses of causation will favor the first alternative. Those who think causation should be cashed out in terms of objective chances will identify dependency hypotheses with compete specifications of conditional chances. And, if one thinks, as I do, that causation is best understood in subjunctive terms, then the Stalnaker/Gibbard/Harper approach will suit one's fancy.[47]

In spite of this ecumenical, irenic attitude, Lewis and Joyce follow Stalnaker, Gibbard, Harper, and Sobel in analysing causal influence in terms of counterfactuals; and Skyrms is happy to accept these counterfactual formulations, though he does take the counterfactuals as unanalysed primitives. Their decision theories were counterfactual, but not so-called in an effort to accommodate reasonable disagreement about the proper analysis of causal influence.

46. Skyrms (1982, p. 697)
47. Joyce (1999, p. 127)

3. Hedden's Objections to Causal Decision Theories

Hedden raises three objections to causal decision theory 'properly so-called'. The only examples of 'properly causal' decision theories he cites are the theories of Lewis (1981) and Skyrms (1980),[48] so they will be my primary focus here. But things will turn out similarly for other versions of causal decision theory. As we saw in §1 above, in certain conditions, there is no difference between the theories of Stalnaker, Gibbard, Harper, Lewis, Joyce, and Sobel. Each of these conditions are met in Hedden's decisions. So, in each of these decisions, all of these theories will agree.

3.1. Overdetermination

Hedden's first objection concerns the case of Overdetermination I introduced in §2.1 above. Hedden acknowledges that every extant version of causal decision theory tells you to not throw in this decision. But he objects that, in so doing, they depart from 'the guiding thought that you should evaluate actions in terms of how likely they are to cause good results'.

It seems to me that there are two potential guiding thoughts, both of them causal in nature. One guiding thought is that, in evaluating an act, you should consider how it affects the world; and you should only care about features of the world over which you might exercise causal control—features which you may be in a position to causally influence. The other guiding thought is that, in evaluating an act, you should consider the act's token effects.

The first guiding thought tells you to not throw, since whether you throw doesn't affect whether the window breaks; you have no causal control over whether the window breaks. The second guiding thought tells you to throw, since if you throw, the window's breaking will be a token effect of your choice. Since it is clear that you have no reason to throw, this decision teaches us that the second guiding thought is mistaken. But it does not speak against the first guiding thought.

As I mentioned in §2.1, while causalists have occasionally used the language of token causation when describing their theories, they have for the most part stuck to the language of causal influence and causal control. So it seems to me that it was the second thought guiding them, and not the first.

3.2. Constitution

Hedden's second objection concerns decisions like this one:

> You have an hour before dinner, during which you can either go for a walk or sit at home. The only things you care about are whether you go

48. Hedden 2023, p. 731

> for a walk and whether you get cake or jello at dinner tonight. You prefer going for a walk, and you prefer getting cake. Dinner has already been prepared, so going for a walk makes no difference to whether you get cake or jello.

The important thing is that, if you go for a walk, this won't just *affect* whether you get what you want. It will also *constitute* you getting what you want. But no act causally influences itself. For this reason, Hedden suggests that causal decision theory will ignore the value of going for a walk. On the other hand, every act counterfactually depends upon itself. So Hedden's counterfatual decision theory will tell you to go for the walk.

In his definition of causal dependence, Lewis explicitly stipulates that, in order for e to causally depend upon c, c and e must be mereologically distinct. For this reason, Lewis requires that the states in the rich partition $\mathcal{S}$ describe matters which are distinct from your choice. Your going for a walk is not mereologically distinct from itself. So your going for a walk does not causally depend upon your going for a walk. So your going for a walk cannot be a *state*.

Does this mean that causal decision theories ignore the value of going for a walk? It does not. Causal decision theories don't calculate utility with the values of states alone. They calculate utility with the values of *outcomes*. And outcomes are act-state conjunctions. So the value of the act itself will be included in the accounting. Let's see how this goes with Lewis's theory. We can model the decision with four worlds, W_{GC}, in which you *g*o for a walk (G) and get *c*ake (C), W_{GJ}, in which you go for the walk and get *j*ello (J), W_{SC}, in which you *s*it at home (S) and get cake, and W_{SJ}, in which you sit at home and get jello. Suppose you spread your probabilities evenly over these four worlds, $P(W_{GC}) = P(W_{GJ}) = P(W_{SC}) = P(W_{SJ}) = 1/4$, and that $V(W_{GC}) = 4, V(W_{GJ}) = V(W_{SC}) = 2$, and $V(W_{SJ}) = 0$.

We can use the state partition $\mathcal{S} = \{C, J\} = \{\{W_{GC}, W_{SC}\}, \{W_{GJ}, W_{SJ}\}\}$. These states describe occurrences (your getting cake and your getting jello, respectively) which are mereologically distinct from how you choose between the options G and S. This state partition will *also* be a partition of causal dependency hypotheses according to Lewis, since your going for a walk or sitting at home won't make any difference to the chances of cake or jello. So the U_2-value of going for a walk, and staying at home, are

$$\begin{aligned} U_2(G) &= P(C) \cdot V(GC) + P(J) \cdot V(GJ) \\ &= 1/2 \cdot 4 + 1/2 \cdot 2 \\ &= 3 \\ \text{and} \quad U_2(S) &= P(C) \cdot V(SC) + P(J) \cdot V(SJ) \\ &= 1/2 \cdot 2 + 1/2 \cdot 0 \\ &= 1 \end{aligned}$$

The value of the act itself is included in expected utility, since each summand includes the value of a conjunction of an act and a causal dependency hypothesis.

Hedden says that Lewis's theory won't tell you to go for the walk because

> For Lewis, dependency hypotheses are maximally specific propositions about how things might be in ways that do not depend causally on your present actions. Offhand, this suggests that facts about which of your present actions is performed could be included in dependency hypotheses. After all, which of your present actions you perform does not *causally* depend on which of your present actions you perform!

If dependency hypotheses are not compatible with each of your available acts, then some conjunction AK would be contradictory, and the value of an act-state conjunction in U_2 would not be well-defined. So neither going for a walk not sitting at home would have a U_2 value, so neither would maximize U_2, and Lewis's theory would fall silent.

Here, Hedden appears to have conflated two different characterisations of causal dependency hypotheses. Skyrms (1980, p. 133) says that they are

> maximally specific specifications of the factors outside our influence at the time of decision.

Lewis, on the other hand, says that causal dependency hypotheses are propositions

> about how the things [the decision maker] cares about do and do not depend causally on his present action.

Neither Skyrms nor Lewis ever characterize dependency hypotheses as 'propositions about how things might be in ways that do not depend causally on your present action". And neither of the characterisations they do give suggest that dependency hypotheses should settle how you choose. For Skyrms's characterisation: how you choose is certainly not outside your influence at the time of decision. For Lewis's: how you choose doesn't say anything about how the things you care about do and do not causally depend upon how you choose. So there's no reason to think this information should be included in a causal dependency hypothesis.

In his informal characterisation of dependency hypotheses, Lewis does seem to presuppose that you only care about states, and you don't care at all about which act you select. Fortunately, this doesn't make any difference to the recommendations of the theory, since the theory allows you to care about which act you select, and it takes that care into account just as it should.

It's worth noting that Lewis explicitly considers the possibility that some act may be incompatible with some dependency hypothesis, raising the worry that '[i]f any of the [probabilities $P(AK)$] is zero, the rule of [U_2]-maximising falls silent'. However, he insists that this possibility will never arise, precisely because dependency hypotheses do not settle how you choose, and so must be compatible with every available act. He tells us that dependency hypotheses

> are causally independent of the agent's actions. They specify his influence over other things, but over them he has no influence...Dependency hypotheses are 'act-independent states' in a causal sense.[49]

He reaffirms this commitment in the postscript added to 'Causal Decision Theory' in his *Philosophical Papers, volume 2*, where he writes that 'I had presupposed...that any option would be compatible with any dependency hypothesis'.[50] This feature of Lewis's theory will also be relevant to Hedden's third objection.

3.3. Determinism

Hedden's third objection concerns this decision:

> You must choose between two bets on the proposition Q. Bet A costs \$10 and pays out \$11 (netting you \$1 profit) iff Q is true. Bet B costs \$1 and pays out \$11 (netting you \$10 profit) iff Q is true.

	Q	$\neg Q$
Bet A	\$1	−\$10
Bet B	\$10	−\$1

> Q is the proposition that the initial state of the universe and the laws together entail that you take bet A. You are all but certain that the laws are deterministic.

This decision is complicated for three reasons. In the first place, there is some ambiguity about the proposition Q. In the second place, we have to be careful about how the payouts of the bets are determined. In the third place, there is controversy about whether the payout of bet A is under your control.

Distinguish two different propositions: Q says that the initial conditions and the laws (whatever they happen to be) entail that you take bet A. $Q_@$ says that the initial conditions and the *actual* laws entail that you take bet A. In $Q_@$, 'the laws' rigidly denote the actual laws, even when we consider other possible worlds. In Q, when we consider other possible worlds, 'the laws' non-rigidly denotes the prevailing laws at those worlds.

Ahmed (2014, p. 124) presents a decision with exactly these payoffs, and in which you are betting on the proposition $Q_@$.[51] But I believe Hedden wishes us to consider

49. Lewis (1981, p. 13)
50. Lewis (1986*b*, p. 337)
51. Or, what comes to the same thing for our purposes, Q_L: the initial conditions and laws L (which you just happen to know are the actual laws) entail that you take bet A.

bets on the proposition Q, for he writes that QB is a metaphysical impossibility. It is metaphysically impossible to take bet B while the initial conditions and the laws entail that you take bet A.[52] But it is not metaphysically impossible to take bet B while the initial conditions and the *actual* laws entail that you take bet A. So I'll assume henceforth that it is the truth of the non-rigid Q which determines the payouts of bets A and B.

Notice that, if QB is impossible, then it doesn't really matter what dollar amount we place in the lower left-hand corner of the table. There is no possibility in which you take bet B and win. If we interpret Q non-rigidly, then, while *others* could win bet B, *you* winning bet B is both epistemically and metaphysically impossible. And causal decision theory pays no mind to impossible outcomes. (This is why I've colored that cell of the table grey.) In Ahmed's decision, in contrast, it was metaphysically (but not epistemically) possible to win bet B. So for that decision, the payout of B mattered.

In the second place, there are some fiddly details about how the payouts of the bets are determined that make a difference to causal decision theory's recommendations. For instance, if we imagine that there's some other flesh-and-blood agent offering you these bets, then it will matter how that agent is determining the truth-value of Q. If they're drawing *inferences* about the truth of Q from your behavior and their knowledge of the deterministic laws, then the payouts of the bets will be under your control, even if the truth of Q is not. If they're basing the payouts on some measurement they've made of the initial conditions, and if the deterministic laws are anything like the laws at our world, then there will be ways of changing the past microstate so that their macromeasurements are unchanged, yet you choose differently.[53] In that case, the truth of Q could come apart from the payouts of the bets, and it could be that the payouts are under your control even if Q is not. Just to bracket all such concerns, suppose the bets are being offered to you by an omniscient deity, and that you'll collect your reward in the afterlife.

In the third place, there is controversy about whether the payout of bet A is under your control. Let's assume that causal control is revealed in causal counterfactuals. Some theories of causal counterfactuals say that the payouts of the bets are not under your control, and others say that they are. There are some who interpret counterfactuals in such a way that they hold fixed every feature of the distant past.[54] For these authors, when we counterfactually suppose that you took a different bet, we imagine that there is a tiny, localized violation of the actual laws just before the moment of choice, and that this violation of the actual laws leads you to take a different bet. And

52. Hedden also says that $\neg QA$ is a metaphysical impossibility; but I don't see how this could be on any interpretation of Q. If you take bet A and the relevant laws are indeterministic, then those laws and the initial conditions needn't entail anything about how you bet. In that case, Q would be false while A is true. I assume he meant that $\neg QA$ is an epistemic impossibility.

53. See Dorr (2016).

54. See Jackson (1977), Lewis (1979), Halpin (1991), Lange (2000), Beebee (2003), and Kment (2006).

there are others who interpret counterfactuals so that they only hold fixed *macroscopic* features of the distant past. For these authors, when we counterfactually suppose that you took a different bet, we imagine that the initial conditions of the universe were ever-so-slightly different on a microscopic scale, and that these microscopic differences only become macroscopically noticeable at the moment of choice.[55] Call the first group 'miracle-lovers', and call the second group 'backtrackers'. Backtrackers about causal counterfactuals hold that, while we have no control over the *macro*state of the past, we exert some causal influence over the past *micro*state. And while this causal influence makes no difference to the vast majority of decisions, it becomes relevant in recherché decisions like this.[56,57]

If we think that causal influence is revealed with causal counterfactuals and we are backtrackers about those counterfactuals, then we will think that you have control over the payout of bet A. If you were to take A, then the initial conditions and the laws together would entail that you take A, so Q would be true, and A would have a payout of \$1. On the other hand, if you take B, then the initial conditions and the laws together would entail that you take B, so Q would be false, and the payout of A would be $-\$10$. So for backtrackers, how much money you get from bet A is fully under your control. While bet B is guaranteed to lose you \$1, taking bet A will make it pay out \$1. In that case, every version of causal decision theory will tell you to take bet A.

If we think causal influence is revealed with counterfactuals and we are miracle-lovers, matters are more complicated. In this decision, there are three relevant epistemic possibilities. It could be that you take bet B, in which case Q is false. It could be that the laws are deterministic and you take bet A, in which case Q is true. Or it could be that the laws are indeterministic and you take bet A. Let's call the first of these possibilities '$W_{\neg QB}$', the second 'W_{QA}', and the third '$W_{\neg QA}$'. The third possibility is *metaphysically* possible, but epistemically negligible, since you are all but certain that you'll take A iff Q is true. Nonetheless, it is relevant insofar as it reveals what causal control you exercise at one of the epistemically possible worlds.

55. See Bennett (1984), Nute (1980, §5.3), Goggans (1992), Albert (2003), Kutach (2002), Loewer (2007), Wilson (2014), Goodman (2015), and Dorr (2016).

56. Not all of these authors are explicit that the counterfactuals they're using are *causal* counterfactuals—the kind of counterfactuals which reveal causal influence and control. And some, like Dorr, are explicit that they are *not* causal counterfactuals. But others are clearly talking about causal counterfactuals. On the miracle-loving side, we've already seen that Lewis thinks that his counterfactuals reveal causal control and that he uses these counterfactuals in his decision theory. On the backtracking side, Loewer (2007) and Albert (2015) are explicit about their counterfactuals revealing agential control; and Loewer calls them 'decision conditionals'.

57. These options aren't exhaustive. As I mentioned in §3 above, my own view is that neither the initial conditions nor the laws of nature causally counterfactually depend upon how you choose; I explain this view in more depth, and explain how it handles different problems for causal decision theory, in Gallow (forthcoming*a*). McNamara (forthcoming) says that, in decisions like these, causal counterfactuals needn't hold fixed *either* the initial conditions *or* the laws. In Hedden's decision, McNamara says that, were to choose A, you'd win \$1, and were you to choose B, you'd lose \$1. So his version of causal decision theory will behave exactly like the backtracker's version of causal decision theory in this decision.

For Lewis's version of causal decision theory, we can take the relevant states to be propositions about how much money you win. That is, we can let $\mathcal{S} = \{-\$1, \$1, -\$10\} = \{W_{\neg QB}, W_{QA}, W_{\neg QA}\}$. (At $W_{\neg QB}$, you lose \$1, at W_{QA}, you get \$1, and at $W_{\neg QA}$, you lose \$10.) These states describe occurrences which are mereologically distinct from how you choose between the bets A and B. Since money is all you care about, we can suppose that $V(\$1) = 1$, $V(-\$1) = -1$, and $V(-\$10) = -10$.

These are the relevant states. The causal dependency hypotheses tell us how these states do and do not causally depend on your choice. For miracle-lovers, there are two epistemically possible dependency hypotheses: if you're at W_{QA}, then you'll get \$1 if you take A and you'll lose \$1 if you take B. Call this dependency hypothesis 'K_A' (since you're all but certain that this dependency hypothesis is true iff you choose A). On the other hand, if you're at $W_{\neg QB}$, then you'll lose \$10 if you take A and you'll lose \$1 if you take B. Call this dependency hypothesis 'K_B' (since you're all but certain that this dependency hypothesis is true iff you choose B).

$$K_A = (A \mathrel{\Box\!\!\rightarrow} \$1) \wedge (B \mathrel{\Box\!\!\rightarrow} -\$1)$$
$$K_B = (A \mathrel{\Box\!\!\rightarrow} -\$10) \wedge (B \mathrel{\Box\!\!\rightarrow} -\$1)$$

At least, it seems to me that these are the two relevant dependency hypotheses for Lewis. But Hedden disagrees for two reasons. In the first place, he suggests that Lewis will take Q and $\neg Q$ to be causal dependency hypotheses. In defense of this, he notes that 'The initial state of the universe and the laws of nature are both beyond your causal control'.[58] Here, Hedden appears to be confusing Skyrms's and Lewis's understandings of causal dependency hypotheses. Skyrms says that causal dependency hypotheses are specifications of factors beyond your control. Lewis says that they are propositions about how the things you care about do and do not causally depend upon your choice. And he says that the dependency hypotheses themselves cannot depend upon how you choose. But Q may depend upon how you choose. In the world W_{QA}, Q is true but would be false, were you to choose B instead, since BQ is metaphysically impossible. So Q cannot be a dependency hypothesis for Lewis. (I'll come back to Skyrms momentarily.)

There is another aspect of the dependency hypothesis K_B which Hedden objects to. He says that, for miracle-lovers like Lewis, the counterfactual $A \mathrel{\Box\!\!\rightarrow} \$1$ is true even if you choose B. If that's right, then Lewis would think that K_A is the only causal dependency hypothesis. To appreciate why Hedden thinks this, consider the world $W_{\neg QB}$ at which you take B, and then counterfactually suppose that you take bet A instead. If you're a miracle-lover, then you will imagine a world—call it '$W_?$' for now—which shares the history of $W_{\neg QB}$, but which contains a localized violation of the actual laws just before the moment of choice; at that time, the actual laws are

58. Hedden (2023, p. 743)

violated in just the way they need to be for you to take bet A.

Hedden suggests that, at $W_?$, Q will be true. That's because he thinks that, at $W_?$, the laws will be both deterministic and unviolated. Of course, the *actual* laws are violated at $W_?$, but that doesn't mean that the laws *at* $W_?$ are violated at $W_?$. If the laws are deterministic and unviolated and you choose A, then the initial conditions and the laws must entail that you choose A. Hence Q must be true. Why does Hedden think that the laws at $W_?$ will be deterministic? In footnote 22, he tells us:

> I think that Lewis' (1973*b*) own best-systems analysis of laws suggests that the laws would still be deterministic [at $W_?$]; the best deductive system—the one which optimally balances simplicity against informational strength—will agree with the actual best deductive system for all regions of spacetime outside the region where the divergence miracle occurs while stipulating exactly what happens within that miraculous region. Then, the laws in [$W_?$] will be a deterministic variant of the actual laws, with a carve-out for the region in which the divergence miracle occurs.

Grant that Hedden has accurately described the best deductive system for $W_?$. Even so, Lewis will not say that the laws of $W_?$ are deterministic. For Lewis, only the *regularities* of the best system get to be laws:

> A law is any regularity that earns inclusion in the ideal system...The ideal system need not consist entirely of regularities; particular facts may gain entry if they contribute enough to collective simplicity and strength...But only the regularities of the system are to count as laws.[59]

This means that—for Lewis—the stipulation of exactly what happens within the miraculous region will not be part of the laws at $W_?$.[60] And so the laws at $W_?$ will not settle what happens in that region. So they will not settle whether you take the bet. So Q will be false. So $W_?$ is a possibility in which Q is false and you take A. Having justified the name, let us now call $W_?$ '$W_{\neg QA}$'. Since $W_{\neg QA}$ is a possibility in which you lose \$10, the counterfactual $A \mathbin{\Box\!\!\rightarrow} -\$10$ will be true at $W_{\neg QB}$. So you are all but certain that the dependency hypothesis K_B will be true if you choose B.

It's worth mentioning one other complication with these dependency hypotheses. It doesn't look like there's any world where you choose B and the dependency hypothesis K_A obtains. After all, we reasoned our way to the conclusion that the epistemically

59. Lewis (1983, p. 367)
60. Albert and Loewer have objected to this part of Lewis's system. They argue we should include the world's low-entropy initial conditions as a law, since this particular fact is needed to derive the second law of thermodynamics. However, even for Albert and Loewer, it's not clear that the stipulation of what happens in the miraculous region would qualify as a law of nature. And, in any case, Albert and Loewer are backtrackers, not miracle-lovers.

possible world in which you choose B is a world in which, had you chosen A instead, there would have been a miracle, and the laws and initial conditions would not have entailed that you choose A. So it seems any B-world is a world in which $A \mathrel{\Box\!\!\rightarrow} \$1$ is false.

This is a potential problem for Lewis's theory, since in order for the term $V(BK_A) = \sum_{W \in \mathcal{W}} P(W \mid BK_A) \cdot V(W)$ to be well-defined, the conjunction BK_A must receive non-zero probability. This is the potential problem we encountered in §3.2 above. As we saw there, it is a problem Lewis was well aware of. And it was very clear how he wanted to deal with it. He writes:

> If any of the [probabilities $P(AK)$] is zero, the rule of [U_2]-maximising falls silent...Should that silence worry us? I think not, for the case ought never to arise...Nothing should ever be held as certain as all that...However much reason you may get to think that option A will not be realized if K holds, you will not if you are rational lower [$P(AK)$] quite to zero. Let it by all means get very, very small; but very, very small denominators do not make utilities go undefined.[61]

So let us imagine a possible world populated by nothing more than trillions upon trillions of duplicates of you, with just your apparent memories, facing a decision just like yours. In this world, you and only you choose B. All the rest choose A. In this world, were you to choose A, the best system would include the generalisation that all who face this decision choose A. So, by Lewis's lights, were you to choose A, it would be a law that you so choose, and Q would be true, winning you \$1. Give this world vanishingly small but non-zero probability, and that will allow us to proceed. You should likewise give some positive but non-zero probability to you choosing B when the dependency hypothesis K_A obtains. Nonetheless, I'll assume that these credences are small enough to be ignored, so that for all practical purposes, you are certain that K_A is true iff you choose A and that K_B is true iff you choose B.

Then, Lewis's theory says that

$$\begin{aligned}
U_2(A) &= P(K_A) \cdot V(AK_A) + P(K_B) \cdot V(AK_B) \\
&= P(A) \cdot V(\$1) + P(B) \cdot V(-\$10) \\
&= P(A) - 10(1 - P(A)) \\
&= 11 \cdot P(A) - 10 \\
\text{and} \quad U_2(B) &= P(K_A) \cdot V(BK_A) + P(K_B) \cdot V(BK_B) \\
&= P(A) \cdot V(-\$1) + P(B) \cdot V(-\$1) \\
&= -1
\end{aligned}$$

61. Lewis (1981, p. 13–14)

where $P(A)$ is your probability that you will choose bet A. If $P(A) > 9/11$, then $U_2(A) > U_2(B)$; whereas, if $P(A) < 9/11$, then $U_2(A) < U_2(B)$. So, for miracle-lovers who think of causal influence in terms of causal counterfactuals, this is a decision in which causal decision theory's advice depends upon which choice you think you'll make. Decisions like these are discussed by Gibbard & Harper (1978), Skyrms (1990), Egan (2007), Arntzenius (2008), and Joyce (2012), among many others. If you begin deliberation thinking that you're most likely to take A, then Lewis's theory will tell you to take A. If you instead begin deliberation thinking that you're most likely to take B, then Lewis's theory will tell you to take B.

You might think that a decision theory shouldn't advise you to take bet B in this decision, even if the payouts of the bets are outside of your control, and even if you're confident that you're going to take B. And you might see this as a reason to worry about causal decision theory. I would urge caution. If you're confident that the bet's payouts are beyond your control, and you're confident that you're going to take B, then you're confident that bet B pays out \$9 more than bet A does. Not only that, but you're confident that nothing you can do would change that fact.

But put that aside. Even if you are worried about what Lewis's theory is saying here, so long as you're a miracle-lover, it's no reason to prefer Hedden's 'counterfactual decision theory'. It's the miraculous counterfactuals which are telling you to take B in this decision, not any other feature of Lewis's theory. So, if you're a miracle-lover, Hedden and Lewis both give the same advice—as does every other form of causal decision theory.

You might think that causal counterfactuals are miraculous, whereas standard English language counterfactuals are backtracking. Then, a decision theory using causal counterfactuals could tell you to take bet B while a version using standard counterfactuals would tell you to take bet A. If you thought that taking bet B was irrational and taking bet A rationally required, that would give you a reason to prefer Hedden's counterfactual decision theory to a version of causal decision theory which used counterfactuals stipulated to hold fixed factors beyond your influence (like the counterfactual formulation of Skyrms's theory). However, Hedden does not argue that standard English counterfactuals backtrack in this decision. His official stance is neutrality about whether the counterfactuals backtrack. He argues that, no matter whether they backtrack, counterfactual decision theory will tell you to take bet A over bet B. (This is almost but not quite right; what's true is that, no matter whether they backtrack, a counterfactual decision theory will tell you to take bet A over bet B *so long as you're confident that you'll take bet A*.)

Why does Hedden think that causal decision theory will disagree with counterfactual decision theory in this decision? He thinks this because he assumes that $\{Q, \neg Q\}$ is a partition of causal dependency hypotheses. And Hedden suggests that, if this is a partition of dependency hypotheses, then causal decision theory will say that A is impermissible, since 'no matter how things beyond your control might be (*i.e.*, no matter whether Q or $\neg Q$ is true), [B] yields a strictly better outcome than [A].'[62] Here, Hedden appeals to a principle of *causal dominance*, which says that an act is impermissible if another act is preferred to it no matter which causal dependency hypothesis obtains.

Dominance B *dominates* A across the partition $\mathfrak{Z}$ iff, for every $Z \in \mathfrak{Z}$, $V(BZ) > V(AZ)$.

Causal Dominance B *causally* dominates A iff B dominates A across a partition of causal dependency hypotheses.

Since causal decision theory says it is impermissible to choose causally dominated acts, Hedden concludes that causal decision theory says A is impermissible

Causal decision theory does forbid causally dominated acts, but even if we grant that $\{Q, \neg Q\}$ is a partition of causal dependency hypotheses, it doesn't follow that B causally dominates A. In order for B to dominate A across the partition $\{Q, \neg Q\}$, we must have $V(BQ) > V(AQ)$. But BQ is metaphysically impossible. So $V(BQ)$ is undefined. So we won't have $V(BQ) > V(AQ)$. So bet B doesn't dominate bet A across the partition $\{Q, \neg Q\}$. Here, the distinction between Hedden's Q and Ahmed's $Q_@$ is important. For miracle-lovers, $\{Q_@, \neg Q_@\}$ *is* a partition of causal dependency hypotheses, and bet $B_@$ *does* dominate bet $A_@$ across this partition. (where '$A_@$' and '$B_@$' are bets just like A and B, except on the proposition $Q_@$.) For, unlike BQ, $B_@Q_@$ is metaphysically possible. So the dominance argument goes through for Ahmed's $Q_@$, but not for Hedden's Q.

(Of course, if $V(BQ)$ is undefined, this makes trouble for the calculation of $U_2(B)$, since one of the summands in $U_2(B)$ is $P(Q) \cdot V(BQ)$. If $V(BQ)$ is undefined, then this product is undefined, so the U_2-value of B will be undefined. However, this is trouble that Lewis will not face, since Lewis denies that $\{Q, \neg Q\}$ is a partition of causal dependency hypotheses. According to Lewis, causal dependency hypotheses must be compatible with each of the available acts. Since Q is not compatible with B, it is not eligible to be a dependency hypothesis.)

Hedden's reason for thinking that $\{Q, \neg Q\}$ is a partition of dependency hypotheses appeals to Skyrms's characterisation. He writes that 'Q is a proposition about the initial state of the universe and the laws of nature [and] the initial state of the universe and the laws of nature are both beyond your causal control.'[63] In order to move from 'the initial conditions are beyond your control' and 'the laws are beyond your control' to 'the initial conditions and the laws are beyond your control', we must appeal to a principle of agglomeration which says that, if X is beyond your control and Y is beyond your control, then XY is beyond your control. This principle is subject to

62. Hedden (2023, p. 743)
63. Hedden (2023, p. 743)

counterexample. For instance: suppose that a 'collapse' interpretation of quantum mechanics is correct and that we have a polarizing filter set to an angle $\cos^{-1}(\sqrt{9/10})$ from the horizontal. This means that a horizontally polarized photon will have a 90% chance of passing through the filter, and a vertically polarized photon will have a 10% chance of passing. You have two options: if you choose option A, then photons x and y will be placed in the correlated entangled state $A = \sqrt{1/2}\ |\uparrow_x, \uparrow_y\rangle + \sqrt{1/2}\ |\rightarrow_x, \rightarrow_y\rangle$ and sent towards the filter. A has a 1/2 chance of collapsing onto a state in which both x and y are polarized vertically, and a 1/2 chance of collapsing onto a state in which both x and y are polarized horizontally. On the other hand, if you choose option B, then photons x and y will be placed in the anti-correlated entangled state $B = \sqrt{1/2}\ |\uparrow_x, \rightarrow_y\rangle + \sqrt{1/2}\ |\rightarrow_x, \uparrow_y\rangle$ and sent towards the filter. B has a 50% chance of collapsing onto a state in which x is polarized vertically and y polarized horizontally, and a 50% chance of collapsing onto a state in which x is polarized horizontally and y polarized vertically. Then, you have no control over whether x passes through the filter. It will have a 50% chance of passing through no matter what you do. Likewise, you have no control over whether y is passed through the filter. It, too, will have a 50% chance of passing through no matter what you do. But you *do* have control over whether *both x and y* pass through the filter. If you choose the correlated state A, then the chance of both x and y passing will be 41%; whereas, if you choose the anti-correlated state B, the chance of both x and y passing will only be 9%.[64]

Perhaps Hedden did not wish to *infer* that you have no control over Q from the fact that you have no control over the laws and you have no control over the initial conditions. Perhaps he simply assumed that you have no control over the conjunction of the laws and the initial conditions, and that, for this reason, you do not have any control over whether Q, and so $\{Q, \neg Q\}$ is a partition of causal dependency hypotheses. As I've already mentioned, Lewis would not accept this claim. Q is incompatible with you choosing bet B, and Lewis insists that a causal dependency hypothesis be compatible with each of your available options.

Skyrms (1980) does not explicitly stipulate that dependency hypotheses are compatible with each of your available acts, but this follows from his characterization of dependency hypotheses together with the following plausible principle: if a factor would change, depending upon which choice you make, then it not outside of your influence. It won't matter how we interpret this counterfactual, so long as it satisfies the principle $(\phi \mathrel{\Box\!\!\rightarrow} \psi) \rightarrow (\Diamond\phi \rightarrow \Diamond\psi)$. Suppose, for *reductio*, that one of Skyrms's $\boldsymbol{k}$'s is incompatible with some available act, A. Then, consider a world at which $\boldsymbol{k}$ is true. Since $\boldsymbol{k}$ is outside of your influence, it would remain true, were you to choose

64. Similar examples appear in Widerker (1987), McKay & Johnson (1996), and Huemer (2000). However, while their counterexamples show that 'you do not have the ability to make X false' and 'you do not have the ability to make Y false' does not imply 'You do not have the ability to make XY false', this counterexample shows that agglomeration fails even for 'you don't have any control over whether'.

A. So at this world, the counterfactual $A \mathrel{\Box\!\!\rightarrow} Ak$ is true. Since A is an available act, it is possible that you choose it. So, by the principle, it is possible that you choose A while k is true. Contradiction. So it seems that Skyrms, too, should accept that dependency hypotheses must be compatible with every available act, so that $\{Q, \neg Q\}$ is not a partition of causal dependency hypotheses.

A. Relations Between Causal Decision Theories

This appendix gives a self-contained explanation of the relationships between the three versions of causal decision theory from §1 which tell you to maximize U_1, U_2, and U_3,

$$U_1(A) = \sum_{O \in \mathcal{O}} P(A \mathrel{\Box\!\!\rightarrow} O) \cdot V(O)$$
$$U_2(A) = \sum_{K \in \mathcal{K}} P(K) \cdot V(AK)$$
$$U_3(A) = \sum_{W \in \mathcal{W}} P^A(W) \cdot V(W)$$

($P^A(X)$ is defined as your expectation of an imaging function, $\sum_{W \in \mathcal{W}} W^A(X) \cdot P(W)$. I will assume throughout that the imaging function satisfies $W^A(A) = 1$, for every $W \in \mathcal{W}$ and every $A \in \mathcal{A}$. It will then follow that $P^A(A) = 1$ and that $P^A(X \mid YA) = P^A(X \mid Y)$.) Many of the results in this appendix are proven elsewhere. I include them here to spare the reader a trek through the literature and to consolidate them with uniform notational conventions.

As a matter of convention, I will say that $x \div 0 = 0$ for any x. Therefore, $P(A \mid B) = 0$ whenever $P(B) = 0$. This will allow us to ignore distracting qualifications about conditional probabilities being defined.

A reminder about notation: $\mathcal{C}h_{\mathcal{X}}$ is a function from worlds to probability distributions over the partition $\mathcal{X}$, with the interpretation that the value of this function given the argument W—which I'll write '$\mathcal{C}h_{\mathcal{X},W}$'—is the chance distribution over the partition $\mathcal{X}$ which obtains at the world W at the moment of choice. And I'll just write '$\mathcal{C}h$' for '$\mathcal{C}h_{\mathcal{W}}$'. If this function is superscripted with an act, $\mathcal{C}h^A_{\mathcal{X}}$, then it will stand for a *counterfactual* chance distribution. This is a function from worlds to probability distributions over $\mathcal{X}$, with the interpretation that the value of this function given the argument W—which I'll write '$\mathcal{C}h^A_{\mathcal{X},W}$'—is the chance distribution that would obtain, were you to choose A. (If there is no chance distribution that would obtain, were you to choose A at W, then this function is undefined.)

Richness of $\mathcal{S}$ A partition of states is *rich* iff the propositions in $\mathcal{S}$ describe occurrences mereologically distinct from your choice and, for each $S \in \mathcal{S}$, each act $A \in \mathcal{A}$, and each dependency hypothesis $K \in \mathcal{K}$, $V(ASK) = V(AS)$.

Outcomes as Act-State Conjunctions Each outcome $O \in \mathcal{O}$ is equivalent to a conjunction AS, for some available act $A \in \mathcal{A}$ and some state $S \in \mathcal{S}$.

Counterfactual Chance Determinacy For each act $A \in \mathcal{A}$ and each world $W \in \mathcal{W}$, there is a chance distribution ch_A such that the counterfactual $A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch_A$ is true at W.

Counterfactual Determinism For each act $A \in \mathcal{A}$ and each world $W \in \mathcal{W}$, there is a state $S_A \in \mathcal{S}$ such that $\mathcal{C}h^A_{\mathcal{S},W}(S_A) = 1$.

Proposition 1. *Assume (i) that there is some rich partition of states $\mathcal{S}$ which makes both counterfactual chance determinacy, counterfactual determinism, and outcomes as act-state conjunctions true, (ii) that $A \mathrel{\square\!\!\rightarrow} \mathscr{C}h_{\mathcal{S}}(S) = 1$ is the same proposition as $A \mathrel{\square\!\!\rightarrow} S$, and (iii) that each dependency hypotheses $K \in \mathscr{K}$ is the conjunction of counterfactuals in a probabilistic full pattern over $\mathcal{S}$,*

$$K = \bigwedge_{A \in \mathscr{A}} A \mathrel{\square\!\!\rightarrow} \mathscr{C}h_{\mathcal{S}} = ch_{A,K}$$

Then, $U_2 = U_1$.

Proof. Given counterfactual chance determinacy, the set of causal dependency hypotheses of the form $\bigwedge_{A \in \mathscr{A}} A \mathrel{\square\!\!\rightarrow} \mathscr{C}h_{\mathcal{S}} = ch_{A,K}$ is a partition. Given counterfactual determinism, every dependency hypothesis is equivalent to a conjunction of the form $K = \bigwedge_{A \in \mathscr{A}} A \mathrel{\square\!\!\rightarrow} \mathscr{C}h_{\mathcal{S}}(S_{A,K}) = 1$. Given assumptions (ii) and (iii), it then follows that each dependency hypothesis has the form $K = \bigwedge_{A \in \mathscr{A}} A \mathrel{\square\!\!\rightarrow} S_{A,K}$.

Since outcomes are act-state conjunctions, $A \mathrel{\square\!\!\rightarrow} O$ is $A \mathrel{\square\!\!\rightarrow} A^*S$. If this counterfactual is possibly true, then $A = A^*$, so each counterfactual $A \mathrel{\square\!\!\rightarrow} O$ with positive probability takes the form $A \mathrel{\square\!\!\rightarrow} AS$, which is equivalent to $A \mathrel{\square\!\!\rightarrow} S$. Since $\mathscr{K}$ is a partition, $A \mathrel{\square\!\!\rightarrow} S$ is equivalent to a disjunction of every dependency hypothesis $K \in \mathscr{K}$ which includes $A \mathrel{\square\!\!\rightarrow} S$ as a conjunct. Call the set of dependency hypotheses including $A \mathrel{\square\!\!\rightarrow} S$ as a conjunct '$\mathscr{K}[A \mathrel{\square\!\!\rightarrow} S]$'. Since distinct dependency hypotheses are incompatible, $P(A \mathrel{\square\!\!\rightarrow} S) = P(\bigvee_{K \in \mathscr{K}[A \mathrel{\square\!\!\rightarrow} S]} K) = \sum_{K \in \mathscr{K}[A \mathrel{\square\!\!\rightarrow} S]} P(K)$.

If $A \mathrel{\square\!\!\rightarrow} S$ is a conjunct of K, then $AK \subseteq S$, by *modus ponens*, and $P(S \mid AK) = 1$. If $A \mathrel{\square\!\!\rightarrow} S$ is not a conjunct of K, then there is *another* state $S^* \in \mathcal{S}$ such that $A \mathrel{\square\!\!\rightarrow} S^*$ is a conjunct of K. In that case $AK \subseteq S^*$, and $P(S \mid AK) = 0$. So $P(S \mid AK) = 1$ if $A \mathrel{\square\!\!\rightarrow} S$ is a conjunct of K, and $P(S \mid AK) = 0$ otherwise. So $\sum_{K \in \mathscr{K}[A \mathrel{\square\!\!\rightarrow} S]} P(K) = \sum_{K \in \mathscr{K}} P(K) \cdot P(S \mid AK)$. And

$$\begin{aligned}
\sum_{O \in \mathscr{O}} P(A \mathrel{\square\!\!\rightarrow} O) \cdot V(O) &= \sum_{S \in \mathcal{S}} P(A \mathrel{\square\!\!\rightarrow} AS) \cdot V(AS) \\
&= \sum_{S \in \mathcal{S}} \left(\sum_{K \in \mathscr{K}} P(K) \cdot P(S \mid AK) \right) \cdot V(AS)
\end{aligned}$$

By the richness of $\mathcal{S}$, $V(AS) = V(ASK)$ for each $S \in \mathcal{S}$ and each $K \in \mathscr{K}$, so

$$\begin{aligned}
\sum_{O \in \mathscr{O}} P(A \mathrel{\square\!\!\rightarrow} O) \cdot V(O) &= \sum_{S \in \mathcal{S}} \left(\sum_{K \in \mathscr{K}} P(K) \cdot P(S \mid AK) \right) \cdot V(ASK) \\
&= \sum_{K \in \mathscr{K}} P(K) \left(\sum_{S \in \mathcal{S}} P(S \mid AK) \cdot V(ASK) \right) \\
&= \sum_{K \in \mathscr{K}} P(K) \cdot V(AK)
\end{aligned}$$

$\square$

Counterfactual Determinacy For every available act $A \in \mathscr{A}$ and every world $W \in$

$\mathcal{W}$, there is an outcome $O_A \in \mathcal{O}$ such that the counterfactual $A \mathbin{\Box\!\!\rightarrow} O_A$ is true at W.

Value-level Outcomes Each outcome settles everything you care about, so that, for each outcome $O \in \mathcal{O}$ and any two worlds $W, W^* \in O$, $V(W) = V(W^*)$.

Corollary 1. *Assume (i) counterfactual determinacy (ii) value-level outcomes, (iii) outcomes as act-state conjunctions, (iv) that $A \mathbin{\Box\!\!\rightarrow} \mathcal{C}h_S(S) = 1$ is the same proposition as $A \mathbin{\Box\!\!\rightarrow} S$; and (v) that each dependency hypothesis $K \in \mathcal{K}$ is a conjunction of counterfactuals specifying which outcome would obtain, were you to choose each act,*

$$K = \bigwedge_{A \in \mathcal{A}} A \mathbin{\Box\!\!\rightarrow} O_{A,K}$$

Then, $U_1 = U_2$.

Proof. By (ii) and (iii), for every $W, W^* \in AS$, $V(W) = V(W^*) \stackrel{\text{def}}{=} \lambda_{AS}$. So

$$\begin{aligned} V(ASK) &= \sum_{W \in \mathcal{W}} P(W \mid ASK) \cdot V(W) = \sum_{W \in AS} P(W \mid ASK) \cdot \lambda_{AS} \\ &= \lambda_{AS} \cdot \sum_{W \in AS} P(W \mid ASK) = \lambda_{AS} \end{aligned}$$

and likewise

$$\begin{aligned} V(AS) &= \sum_{W \in \mathcal{W}} P(W \mid AS) \cdot V(W) = \sum_{W \in AS} P(W \mid AS) \cdot \lambda_{AS} \\ &= \lambda_{AS} \sum_{W \in AS} P(W \mid AS) = \lambda_{AS} \end{aligned}$$

So $V(AS) = V(ASK)$, and the partition $\mathcal{S}$ is rich.

(i) and (iv) imply counterfactual chance determinacy, and (v) implies that each dependency hypothesis is a conjunction of counterfactuals in a probabilistic full pattern over $\mathcal{S}$.

So (i), (ii), (iii), (iv), and (v) together imply the assumptions of proposition 1. $\square$

Definition 1. A partition $\mathcal{K}$ is *chance-determining* iff, for every $A \in \mathcal{A}$ and each $K \in \mathcal{K}$, $AK \neq \varnothing$, and there is a chance function $ch_{A,K}$ so that $\mathcal{C}h = ch_{A,K}$ is true at every world in AK.

Definition 2. A chance-determining partition $\mathcal{K}$ is *admissible* iff, for every $X \subseteq \mathcal{W}$, $A \in \mathcal{A}$, and $K \in \mathcal{K}$, $P(X \mid AK) = ch_{A,K}(X)$.

Proposition 2. *Let $\mathcal{K}$ be an admissible chance-determining partition. And for each $K \in \mathcal{K}$, each $W \in K$, and each $A \in \mathcal{A}$, let $W^A = ch_{A,K}$. Then, $U_3 = U_2$.*

Proof. Because $\mathcal{K}$ is admissible, for each $K \in \mathcal{K}$ and $A \in \mathcal{A}$, $P(AK \mid AK) = 1 = ch_{A,K}(AK)$. And, for any $K^* \neq K$ or $A^* \neq A$, $P(A^*K^* \mid AK) = 0 = ch_{A,K}(A^*K^*)$. So every potential chance function $ch_{A,K}$ is certain of the act and $\mathcal{K}$-cell that determine it. So, for each $K \in \mathcal{K}$ and each $W \in K$, $W^A(K) = ch_{A,K}(K) = 1$, and if $K^* \neq K$, $W^A(K^*) = 0$. So

$$P^A(K) = \sum_{W \in \mathcal{W}} W^A(K) \cdot P(W) = \left(\sum_{W \in K} 1 \cdot P(W) \right) + \left(\sum_{W \notin K} 0 \cdot P(W) \right) = P(K)$$

For any $W \in K$,

$$\begin{aligned} P^A(W \mid K) &= \frac{P^A(W)}{P^A(K)} = \frac{\sum_{W^* \in \mathcal{W}} W^{*A}(W) \cdot P(W^*)}{P(K)} = \frac{\sum_{W^* \in K} ch_{A,K}(W) \cdot P(W^*)}{P(K)} \\ &= \frac{ch_{A,K}(W) \cdot \sum_{W^* \in K} P(W^*)}{P(K)} = \frac{ch_{A,K}(W) \cdot P(K)}{P(K)} = ch_{A,K}(W) \end{aligned}$$

And $P(W \mid AK) = ch_{A,K}(W)$, by the admissibility of $\mathcal{K}$. So, for any $W \in K$, $P^A(W \mid K) = P(W \mid AK)$.

Therefore,

$$\begin{aligned} U_2(A) &= \sum_{K \in \mathcal{K}} P(K) \cdot \sum_{W \in \mathcal{W}} P(W \mid AK) \cdot V(W) \\ &= \sum_{K \in \mathcal{K}} \sum_{W \in K} P^A(K) \cdot P^A(W \mid K) \cdot V(W) \\ &= \sum_{W \in \mathcal{W}} P^A(W) \cdot V(W) \\ &= U_3(A) \end{aligned}$$

□

Images as Counterfactual Chances For each world $W \in \mathcal{W}$ and each act $A \in \mathcal{A}$, $W^A = \mathcal{C}h_W^A$.

Corollary 2. *Assume (i) counterfactual chance determinacy, (ii) that images are counterfactual chances, (iii) that each dependency hypothesis $K \in \mathcal{K}$ is a conjunction of the counterfactuals in a probabilistic full pattern over $\mathcal{S}$,*

$$K = \bigwedge_{A \in \mathcal{A}} A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch_{A,K}$$

and (iv) that, for each dependency hypothesis $K \in \mathcal{K}$, $P(X \mid AK) = ch_{A,K}(X)$ for each $X \subseteq \mathcal{W}$.

Then, $U_3 = U_2$.

Proof. By counterfactual chance determinacy, some dependency hypothesis is true at every world. And by conditional non-contradiction, no two dependency hypotheses are true at the same world, else there'd be two true counterfactuals, $A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch$ and $A \mathrel{\Box\!\!\rightarrow} \mathcal{C}h_{\mathcal{S}} = ch'$, where $ch \neq ch'$. So the set of dependency hypotheses form a partition.

By counterfactual chance determinacy, $\mathcal{K}$ is a chance determining partition, and $\mathcal{K}$ is admissible by assumption. By definition, $ch_{A,K} = \mathcal{C}h_W^A$, for any $W \in K$. By images as counterfactual chances, $W^A = \mathscr{C}h^A_W$. So $W^A = ch_{A,K}$. So the assumptions of proposition 2 are satisfied, and $U_3 = U_2$. $\square$

Definition 3. One partition, $\mathscr{k}$, is a *refinement* of another, $\mathscr{K}$, iff, for every $K \in \mathscr{K}$, there's some $k_1, k_2, \ldots, k_n \in \mathscr{k}$ such that $K = \bigcup_{i=1}^n k_i$.

Proposition 3. *Let $\mathscr{K}$ and $\mathscr{k}$ be any two admissible chance-determining partitions such that $\mathscr{k}$ is a refinement of $\mathscr{K}$. Then,*

$$\sum_{K \in \mathscr{K}} P(K) \cdot V(AK) = \sum_{k \in \mathscr{k}} P(k) \cdot V(Ak)$$

Proof. First note that, if $k \subseteq K$, then $Ak \subseteq AK$, so $ch_{A,k} = ch_{A,K}$. So

$$\begin{aligned} V(AK) &= \sum_{W \in \mathscr{W}} P(W \mid AK) \cdot V(W) = \sum_{W \in \mathscr{W}} ch_{A,K}(W) \cdot V(W) \\ &= \sum_{W \in \mathscr{W}} ch_{A,k}(W) \cdot V(W) = \sum_{W \in \mathscr{W}} P(W \mid Ak) \cdot V(W) = V(Ak) \end{aligned}$$

So

$$\begin{aligned} \sum_{k \in \mathscr{k}} P(k) \cdot V(Ak) &= \sum_{K \in \mathscr{K}} \sum_{\substack{k \in \mathscr{k}: \\ k \subseteq K}} P(k) \cdot V(Ak) = \sum_{K \in \mathscr{K}} \sum_{\substack{k \in \mathscr{k}: \\ k \subseteq K}} P(k) \cdot V(AK) \\ &= \sum_{K \in \mathscr{K}} V(AK) \sum_{\substack{k \in \mathscr{k} \\ k \subseteq K}} P(k) = \sum_{K \in \mathscr{K}} V(AK) \cdot P(K) \end{aligned}$$

$\square$

Definition 4.

$$V^Y(X) \stackrel{\text{def}}{=} \sum_{W \in \mathscr{W}} P^Y(W \mid X) \cdot V(W)$$

Lemma 1 (Joyce, 1999, §5.5). *For any partition $\mathfrak{X}$, $V^Y(X) = \sum_{Z \in \mathfrak{X}} P^Y(Z \mid X) \cdot V^Y(XZ)$*

Proof.

$$\begin{aligned} \sum_{Z \in \mathfrak{X}} P^Y(Z \mid X) \cdot V^Y(XZ) &= \sum_{Z \in \mathfrak{X}} P^Y(Z \mid X) \cdot \sum_{W \in \mathscr{W}} P^Y(W \mid XZ) \cdot V(W) \\ &= \sum_{Z \in \mathfrak{X}} P^Y(Z \mid X) \cdot \sum_{W \in Z} P^Y(W \mid XZ) \cdot V(W) \\ &= \sum_{Z \in \mathfrak{X}} \sum_{W \in Z} P^Y(W \mid XZ) \cdot P^Y(Z \mid X) \cdot V(W) \\ &= \sum_{W \in \mathscr{W}} P^Y(W \mid X) \cdot V(W) \\ &= V^Y(X) \end{aligned}$$

$\square$

Corollary 3. *For any partition $\mathfrak{X}$, $U_3(A) = V^A(A) = \sum_{Z \in \mathfrak{X}} P^A(Z) \cdot V^A(Z)$.*

Proof. Immediate from lemma 1, the definition of V^A, and the fact that $P^A(A) = 1$. □

Lemma 2. *If outcomes are value-level, then, for any X compatible with the outcome O, $V^A(OX) = V(OX)$.*

Proof. If outcomes are value-level, then for any $O \in \mathcal{O}$, any X compatible with O, and any $W, W^* \in OX$, $V(W) = V(W^*) = \lambda$, so

$$V^A(OX) = \sum_{W \in \mathcal{W}} P^A(W \mid OX) \cdot V(W) = \sum_{W \in O} P^A(W \mid OX) \cdot \lambda = \lambda \cdot \sum_{W \in O} P^A(W \mid OX) = \lambda$$

And

$$V(OX) = \sum_{W \in \mathcal{W}} P(W \mid OX) \cdot V(W) = \sum_{W \in O} P(W \mid OX) \cdot \lambda = \lambda \cdot \sum_{W \in O} P(W \mid OX) = \lambda$$

□

Definition 5. $U_4(A) \stackrel{\text{def}}{=} \sum_{O \in \mathcal{O}} P^A(O) \cdot V(O)$.

Corollary 4. *If outcomes are value-level, then $U_3 = U_4$.*

Proof. Immediate from corollary 3 and lemma 2. □

Proposition 4. *Assume (i) that outcomes are act-state conjunctions and value-level (ii) that images are counterfactual chances, (iii) that $A \,\Box\!\!\rightarrow \mathcal{C}h(S) = 1$ is the same proposition as $A \,\Box\!\!\rightarrow S$, and (iv) counterfactual determinacy. Then, $U_3 = U_1$.*

Proof. By (ii) and (iii), $A \,\Box\!\!\rightarrow S = \{W \in \mathcal{W} \mid W^A(S) = 1\}$. By (ii), (iii), and (iv), for each $A \in \mathcal{A}$ and each $W \in \mathcal{W}$, there is a $W^* \in \mathcal{W}$ such that $W^A(W^*) = 1$ (that is: images are sharp). So, for each W, either $W^A(S) = 1$ or $W^A(S) = 0$. If $W^A(S) = 1$, then $W \in A \,\Box\!\!\rightarrow S$, and if $W^A(S) = 0$, then $W \notin A \,\Box\!\!\rightarrow S$. So

$$P(A \,\Box\!\!\rightarrow S) = \sum_{W \in \mathcal{W}} W^A(S) \cdot P(W) = P^A(S)$$

By (i) and lemma 2, $V^A(S) = V^A(AS) = V(AS) = V(O)$. So, by corollary 3,

$$\begin{aligned} U_1(A) &= \sum_{O \in \mathcal{O}} P(A \,\Box\!\!\rightarrow O) \cdot V(O) = \sum_{S \in \mathcal{S}} P(A \,\Box\!\!\rightarrow S) \cdot V(AS) \\ &= \sum_{S \in \mathcal{S}} P^A(S) \cdot V^A(S) = U_3(A) \end{aligned}$$

□

References

Ahmed, Arif. 2014. *Evidence, Decision and Causality*. Cambridge, UK: Cambridge University Press.

Albert, David Z. 2003. *Time and Chance*. Cambridge, MA: Harvard University Press.

Albert, David Z. 2015. *After Physics*. Cambridge, Massachusetts: Harvard University Press.

Arntzenius, Frank. 2008. "No regrets, or: Edith Piaf revamps decision theory." In *Erkenntnis*, **68** (2): 277–297.

Beebee, Helen. 2003. "Local miracle compatibilism." In *Noûs*, **37** (2): 258–277.

Bennett, Jonathan. 1984. "Counterfactuals and Temporal Direction." In *The Philosophical Review*, **93** (1): 57–91.

Dorr, Cian. 2016. "Against Counterfactual Miracles." In *The Philosophical Review*, **125** (2): 241–286.

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

Gallow, J. Dmitri. forthcoming*a*. "Causal Counterfactuals without Miracles or Backtracking." In *Philosophy and Phenomenological Research*.

Gallow, J. Dmitri. forthcoming*b*. "Decision and Foreknowledge." In *Noûs*.

Gibbard, Allan & Harper, William L. 1978. "Counterfactuals and Two Kinds of Expected Utility." In *Foundations and Applications of Decision Theory*, edited by A. Hooker, J.J. Leach, & E.F. McClennan, Dordrecht: D. Reidel, 125–162.

Goggans, Phillip. 1992. "Do the Closest Counterfactual Worlds Contain Miracles?" In *Pacific Philosophical Quarterly*, **73** (2): 137–149.

Goodman, Jeremy. 2015. "Knowledge, Counterfactuals, and Determinism." In *Philosophical Studies*, **172** (9): 2275–2278.

Halpin, John F. 1991. "The miraculous conception of counterfactuals." In *Philosophical Studies*, **63** (3): 271–290.

Hedden, Brian. 2023. "Counterfactual Decision Theory." In *Mind*, **132** (527). doi: 10.1093/mind/fzac060.

Hitchcock, Christopher. 2013. "What is the 'Cause' in Causal Decision Theory?" In *Erkenntnis*. DOI 10.1007/s10670-013-9440-9.

Hitchcock, Christopher. 2016. "Conditioning, Intervening, and Decision." In *Synthese*, **194** (4): 1157–1176.

Horgan, Terence. 1981. "Counterfactuals and Newcomb's Problem." In *The Journal of Philosophy*, **78** (6): 331–356.

Huemer, Michael. 2000. "Van Inwagen's Consequence Argument." In *The Philosophical Review*, **109** (4): 525–544.

Jackson, Frank. 1977. "A causal theory of counterfactuals." In *Australasian Journal of Philosophy*, **55** (1): 3–21.

Joyce, James M. 1999. *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Joyce, James M. 2012. "Regret and instability in causal decision theory." In *Synthese*, **187** (1): 123–145.

Kment, Boris. 2006. "Counterfactuals and Explanation." In *Mind*, **115**: 261–309.

Kutach, Douglas. 2002. "The entropy theory of counterfactuals." In *Philosophy of Science*, **69** (1): 82–104.

Lange, Marc. 2000. *Natural Laws in Scientific Practice*. Oxford: Oxford University Press.

Lewis, David K. 1973*a*. "Causation." In *The Journal of Philosophy*, **70** (17): 556–567.

Lewis, David K. 1973*b*. *Counterfactuals*. Malden, MA: Blackwell Publishers.

Lewis, David K. 1979. "Counterfactual Dependence and Time's Arrow." In *Noûs*, **13** (4): 455–476.

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

Lewis, David K. 1983. "New Work for a Theory of Universals." In *Australasian Journal of Philosophy*, **61** (4): 343–377.

Lewis, David K. 1986*a*. "Events." In *Philosophical Papers*, New York: Oxford University Press, volume II, 241–269.

Lewis, David K. 1986*b*. "Postscript to 'Causal Decision Theory'." In *Philosophical Papers, volume 2*, Oxford: Oxford University Press.

Lewis, David K. 1986*c*. "Postscripts to 'Causation'." In *Philosophical Papers*, New York: Oxford University Press, volume II.

Loewer, Barry. 2007. "Counterfactuals and the Second Law." In *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, edited by Huw Price & Richard Corry, Oxford: Oxford University Press, chapter 11, 293–326.

McKay, Thomas J. & Johnson, David. 1996. "A Reconsideration of an Argument against Compatibilism." In *Philosophical Topics*, **24** (2): 113–122.

McNamara, Calum. forthcoming. "Causal Decision Theory, Context, and Determinism." In *Philosophy and Phenomenological Research*. doi:10.1111/phpr.13021.

Meek, Christopher & Glymour, Clark. 1994. "Conditioning and Intervening." In *The British Journal for the Philosophy of Science*, **45**: 1001–1021.

Nozick, Robert. 1969. "Newcomb's Problem and Two Principles of Choice." In *Essays in Honor of Carl G. Hempel*, edited by Nicholas Rescher, Dordrecht: D. Reidel, 114–146.

Nute, Donald. 1980. *Topics in Conditional Logic*. Dordrecht: D. Reidel.

Skyrms, Brian. 1980. *Causal Necessity*. New Haven: Yale University Press.

Skyrms, Brian. 1982. "Causal Decision Theory." In *Journal of Philosophy*, **79** (11): 695–711.

Skyrms, Brian. 1984. *Pragmatics and Empiricism*. New Haven: Yale University Press.

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

Sobel, Jordan Howard. 1986. "Notes on Decision Theory: Old Wine in New Bottles." In *Australasian Journal of Philosophy*, **64** (4): 407–437. doi:10.1080/00048408612342621.

Sobel, Jordan Howard. 1994. *Taking Chances: Essays on Rational Choice*. Cambridge: Cambridge University Press.

Stalnaker, Robert C. 1980. "A Defense of Conditional Excluded Middle." In *Ifs*, edited by W. L. Harper, R. Stalnaker, & G. Pearce, Dordrecht: D. Reidel, 87–104.

Stalnaker, Robert C. 1981. "Letter to David Lewis." In *Ifs*, edited by William Harper, Robert Stalnaker, & Glenn Pearce, Dordrecht: D. Reidel Publishing Company, 151–152.

Stern, Reuben. 2017. "Interventionist Decision Theory." In *Synthese*, **194**: 4133–4153.

Stern, Reuben. 2018. "Diagnosing Newcomb's Problem with Causal Graphs." In *Newcomb's Problem*, edited by Arif Ahmed, Cambridge: Cambridge University Press, chapter 10, 201–220.

Widerker, David. 1987. "On an Argument for Incompatibilism." In *Analysis*, **47** (1): 37–41.

Wilson, Jessica M. 2014. "Hume's Dictum and the asymmetry of counterfactual dependence." In *Chance and Temporal Asymmetry*, edited by Alastair Wilson, Oxford: Oxford University Press, 258–279.