Conjoined Cases

The spectre of isomorphs haunts counterfactual theories of causation, and the standard attempt to dispel it involves employing normality evaluations. This attempt doesn't work, I claim.

Counterfactual theories of deterministic causation (Beckers & Vennekens 2018, Glymour & Wimberly 2007, Hitchcock 2001, Woodward, 2003, Halpern & Pearl 2005, Weslake 2015) analyze causal claims in terms of counterfactuals (§1). According to a typical counterfactual theory, one event is a cause of another in a target situation *iff* the latter counterfactually depends on the former in a witness—some possible situation, carefully constructed from the target situation. The theories typically differ on what witnesses are allowed.

Counterfactual theories face the problem of isomorphs (§2): two cases, called *isomorphs*, share the counterfactual structure although only one case exemplifies causation (Hall 2007, Hiddleston 2005). As counterfactual theories evaluate causal claims based solely on the case's counterfactual structure, they are bound to misjudge one of the cases. To break the isomorphism, some authors have incorporated normality evaluations into their theories (Hall 2007, Hitchcock 2007, Halpern & Hitchcock 2015, Halpern 2016, Gallow 2020). Specifically, these (call them) *normality theories* ask you to assign normality to actual events and their possible alternatives, and you can, in a principled way, do it differently in either isomorph. The theories allow you to construct a witness only out of only sufficiently normal alternatives to actual events. Hence, the theories can distinguish between isomorophs despite the shared counterfactual structure.

Ultimately, as I argue, this strategy fails (§3). Classic pairs of isomorphs include: overdetermination and bogus prevention, causal and non-causal omissions, a cause and background condition, and early preemption and a short circuit. From every such pair, I construct a new *conjoined case* so that however you assign normality to the events, normality theories misjudge one of the causal claims true of the case. Crucially, the judgments lost are the ones that normality evaluations were supposed to save.

Subsequently (§4), I respond to possible objections: that I ought not to conjoin cases in the first place, that my models of the cases are inapt, that the conjoined cases can be handled by an alternative normality theory, that they can be handled if we focus on the normality of entire witnesses rather than events and their alternatives, and that my criticism is old news.

1 Counterfactual theories

The standard measure of success of theories of causation is how well they predict causal judgments that they cannot convincingly explain away (Paul & Hall 2013). This is why much energy in this literature has been spent on formulating cases that support one theory over its competitors. With respect to this measure, counterfactual theories (Beckers & Vennekens 2018, Glymour & Wimberly 2007, Hitchcock 2001, Woodward, 2003, Halpern & Pearl 2005, Weslake 2015) and subsequent normality theories (Hall 2007, Hitchcock 2007, Halpern & Hitchcock 2015, Halpern 2016, Gallow 2020) have been most successful.

Counterfactual theories are formulated within the framework of causal models (Pearl 2009, Spirtes et al. 2001). A *causal model* consists of variables, their ranges, and equations. Variable values denote atomic events, and any non-atomic event is a Boolean combination of atomic events. The value of every variable is determined by an equation from the values of other variables, called the target variable's *parents*. An exogenous equation has no arguments and thus ascribes to the variable its actual value. An endogenous equation has positive arity and determines what happens from what has already happened.[1] An *ancestor* of a variable is a parent of that variable or an ancestor of some parent of that variable. An *assignment* is a function that assigns to every variable a value from the variable's range. A *solution* to a model is an assignment that satisfies the model's structural equations. I'll visualize models using diagrams: nodes correspond to variables (and hence I'll use these terms interchangeably), values in nodes denote the variables' actual values, and edges go from parents to the nodes they parent. I'll deal with acyclic models only, i.e., ones where no node is its own ancestor. A graph of an acyclic model is acyclic (whence the name)–you can't walk along the edges from a node and return to this very node. Acyclic models always have a single solution.

For an illustration, take *overdetermination*:

[1] The revolutionary and the spy don't know about each other. It's high time, either thinks.

[1]Endogenous equations themselves encode what's called *primitive non-backtracking would-counterfactuals* (Hitchcock 2001, Lewis 1979, Woodward 2003). Roughly, evaluating non-backtracking counterfactuals, you think about how the event in the antecedent would affect the future ("if the Fourth Crusade hadn't been organized, Constantinople wouldn't have fallen"). Evaluating backtracking counterfactuals, you think about how the past would have to be for the antecedent to have happen ("if Constantinople hadn't had fallen, it would mean the Fourth Crusade had not been organized"). It hence makes sense to understand exogenous equations as encoding primitive modals: $X \leftarrow x$ states that $X = x$ is necessary in the model–$X = x$ will/would happen in the situation unless a change comes from without–for on deterministic models, necessary is the same as actual. This interpretation of exogenous equations becomes obvious once you allow also for primitive possibility claims (redacted).

They poison the tsar's tea. He dies.

a. The revolutionary poisoning the tea caused the tsar to die.

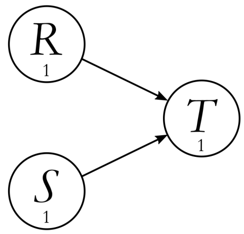

Figure 1. Overdetermination.

The standard representation of overdetermination is this. As atomic events take: $R = 1$ if the revolutionary poisons the tsar's cup, $R = 0$ if he doesn't; $S = 1$ if the spy poisons the tsar's cup, $S = 0$ if she doesn't; $T = 1$ if the tsar dies, $T = 0$ if he survives (fig. 1). Three equations determine what happens:

$$R \leftarrow 1, \quad S \leftarrow 1, \quad T \leftarrow R \vee S \tag{1}$$

The first equation states that the revolutionary poisons the cup. The second, that the spy poisons the cup too. The third, that the tsar dies if the cup has been poisoned at least once. The equations yield a solution: $R = S = T = 1$.

The *counterfactual* "had $X_1 = x_1 \wedge \ldots \wedge X_n = x_n$ happened, φ would have happened," where $X_i = x_i$ are atomic events, and φ is a (possibly non-atomic) event, holds in a model *iff* φ is satisfied by the solution to the model where every X_i's equation is replaced with $X_i \leftarrow x_i$. For instance, "had the spy failed to poison the cup ($S = 0$), the tsar still would have died ($T = 1$)" holds because $T = 1$ after replacing S's equation with $S \leftarrow 0$ in (1): the revolutionary's revolutionary act suffices to kill the tsar.

Counterfactuals underlie *counterfactual theories* of causation (Beckers & Vennekens 2018, Halpern & Pearl 2005, Hitchcock 2001, Weslake 2015, Woodward 2003). For focus, I'll work with Hitchcock's (2001); there are more successful theories out there, but his is simple enough while working well with the examples used here. On his theory, you need to choose a path from the cause variable C and the effect variable E. $C = c$ causes $E = e$ in a model *iff* there are *contrast cause* $C = \underline{c}$ and *contingency* $T_1 = t_1 \wedge \ldots \wedge T_n = t_n$, where *contingency variables* T_i are off-path, such that

act $C = c$ and $E = e$ happened,

wit had $C = \underline{c} \wedge T_1 = t_1 \wedge \ldots \wedge T_n = t_n$ happened, $E = e$ wouldn't have happened, *adm* had $C = c \wedge T_1 = t_1 \wedge \ldots \wedge T_n = t_n$ happened, all the actual events on the chosen path from C to E would have happened (including $E = e$).

Wit requires that the effect doesn't happen under the contrast cause and contingency. Call the solution used in this condition the *witness* of the effect's counterfactual dependence on the cause under the contingency.[2] *Adm* requires that the contingency is admissible: if it had happened together with the cause, then every actual event on the chosen path from the cause to the effect (including the effect itself) still would have happened. The role of the contingency is to unmask the causal process from the cause to the effect. Couched in these terms, *adm* demands that the contingency leaves intact the causal process from the cause to the effect along the chosen path.

The theory handles [1]: the revolutionary poisoning the cup ($R = 1$) caused the tsar to die ($T = 1$). Choose the revolutionary failing to act ($R = 0$) as the contrast cause and the spy failing to act ($S = 0$) as the contingency. Per *act*, $R = 1$ and $T = 1$ happened. Per *wit*, had neither poisoned the cup, $S = 0 \wedge R = 0$, the tsar would have survived, $T = 0$. Per *adm*, $S = 0$ is an admissible contingency because the tsar still would have died if the spy alone had poisoned the cup.

2 Isomorphism and normality

Any counterfactual theory faces the problem of isomorphs (Hall 2007, Hiddleston 2005): there are cases that share the counterfactual structure—i.e., they are naturally represented with the same equations—but that elicit different intuitions. Since a counterfactual theory yields a causal judgment based only on the model of the case, no such theory can handle isomorphs.

Consider *bogus prevention* (Hiddleston 2005):

[2] The spy has a change of mind: she won't poison the tsarina's cup. The tsarina has a good guard, very careful. The guard adds an antidote to the tsarina's harmless tea. She survives.

a. The guard adding the antidote to the cup *doesn't* cause the tsarina to survive.

For reasons that should become clear momentarily, I'll represent the events rather inconveniently: $R = 1$ if the guard adds the antidote to the cup, $R = 0$ if not; $S = 1$ if the spy does *not* poison the cup, $S = 0$ if she does; $T = 1$ if the tsarina survives, $T = 0$ if she dies. Three counterfactuals

[2] I.e., the witness is the solution to the model produced by replacing C's and T_i's equations with $C \leftarrow \underset{\sim}{c}$ and $T_i \leftarrow t_i$.

represent the plot:

$$R \leftarrow 1, \quad S \leftarrow 1, \quad T \leftarrow R \vee S. \tag{2}$$

The last equation states that the tsarina will survive if the guard adds the antidote or the spy decides not to poison the cup.

Never mind what any counterfactual theory says about this case—just compare the model with the one for overdetermination (fig. 1). Their variables, their ranges, and their equations agree. Yet, the revolutionary lacing the cup ($R = 1$) *does* cause the tsar to die ($T = 1$) in [1], while the guard adding the antidote to the tea (again, $R = 1$) *doesn't* cause the tsarina to survive (again, $T = 1$) in [2]. You have isomorphic models but opposing judgments. Thus, relying on structural equations alone cannot do justice to both cases—provided they are correctly represented by the models.

To save counterfactual theories, you can deny that equations (2) do justice to the counterfactual structure of the case. In this vein, Blanchard and Schaffer (2017) claim the model misrepresents [2] because a crucial variable is omitted: one for whether the antidote counteracts the poison. Add the variable, and the judgment is accounted for (fig. 2).

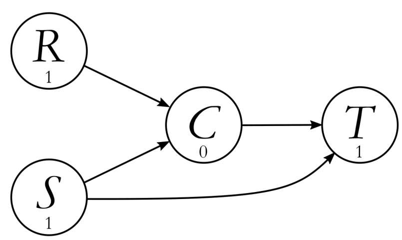

Figure 2. Bogus prevention with an extra node.

Let $C = 1$ mean that the guard's antidote counteracts the poison, and $C = 0$ that it doesn't. Since the antidote neutralizes the poison only if the guard adds the antidote while there is poison to neutralize, and the tsarina survives if the cup is replaced or the poison neutralized, the equations read:

$$R \leftarrow 1, \quad S \leftarrow 1, \quad C \leftarrow R \wedge \neg S, \quad T \leftarrow C \vee S. \tag{3}$$

The sole path from R to T goes through C, and therefore S is the only potential contingency variable. You can't freeze S at 0, for that would entail $C = 1$, contra *adm*. But if $S = 1$, T doesn't counterfactually depend on R. *Act-wit-adm* correctly predict that the eager guard didn't cause the tsarina to survive.

Another solution—the main target of my argument—is to extend counterfactual theories with normality evaluations (Hall 2007, Halpern 2016, Halpern & Hitchcock 2015, Menzies 2017). Implementations differ, but almost all such theories stem from the same principle: the witness must be at least as normal as the actual solution (where the opposite of 'at least as normal' is 'less normal or incomparable'), and the normality of a witness depends (solely or at least significantly) on the normality of the events that constitute the witness.

The concept of normality, as applied to events, is supposed to be (or at least to explicate) the everyday concept of normality. For McGrath (2005:138, see also Thomson 2003:100), who uses the concept to investigates causal omissions, it's normal for an object to φ *iff* the object is supposed to φ. What an object is supposed to do, says McGrath, depends on the category it belongs to: artifacts have intended functions, organs have normal functions, organisms have natural behaviors, moral agents have norms to follow, and so on. Halpern and Hitchcock (2010:402, 2015:431-2) concur. The normality of an event, they say, depends on the frequency of the type of the event in question, moral and social norms, and norms of proper functioning. Other speak of default (rather than normal) events—the events that would have transpired, had noting affected the system from without (Hall 2007, Maudlin 2004, Menzies 2004, Wolff 2016). But in any normality theory, normality is a measure ascribed to variable values and subsequently to entire assignments of values to the model's variables.[3]

To incorporate the distinction into Hitchcock's theory, add a fourth condition:

nrm the witness from *wit* is at least as normal as the solution to the original model.

The condition requires comparing value assignments—specifically, solutions—and authors differ on how that should be done.[4] Typically, they order a variable's values from the most to least normal and subsequently derive the order on assignments. So, Hall (2007) deems one assignment at least as normal as another *iff* every value on the former is at least as normal as the corresponding value on the latter. Halpern and Hitchcock (2015:435) espouse this view too but with two caveats: the outcome of an endogenous structural equation is prima facie normal,[5] and if you can't compare assignments in virtue of values alone, you can invoke special considerations to make the comparison (e.g., one situation happens more often and thus is more normal than the other). I'll call counterfactual theories extended with normality evaluations *normality theories* and use *act-nrm* as their representative; when applied to the simple cases I raise, the four conditions mimic the workings of the more

[3]Halpern and Hitchcock (2015) call events (ab)normal and witnesses (a)typical. I'll just use one term.

[4]The solution to a model is the assignment of values that satisfies the model's equations. A witness is also an assignment, for it's a solution to the post-intervention model specified by *wit*.

[5]This is a nod toward the notion of defaults, for an endogenous equation says what would have happened sans an intervention, i.e., sans perturbing the system.

sophisticated theories of Hall's (2007), Halpern and Hitchcock's (2015), or Menzies's (2017).

Normality theories can distinguish between overdetermination and bogus prevention. In [1], assume that an action (here, poisoning the cup) is less normal than an inaction (here, refraining from poisoning the cup). Therefore, the assignment where neither actor poisons the cup ($R = S = 0$) is more normal than the actual solution, where both do ($R = S = 1$). Hence, per *nrm*, this assignment makes for a good witness. Since the remaining conditions still hold, the theory yields that the revolutionary's act caused the tsar to die. Assume the same for [2]: an action is less normal than an inaction, and hence poisoning the cup is less normal than not doing so. But that means that an assignment where the spy poisons the cup, $S = 0$, is less normal than one where she doesn't, $S = 1$. Therefore, per *nrm*, you can use only the latter as a witness. But on such a witness, whether the tsarina survives doesn't counterfactually depend on whether the antidote is in the cup. The theory correctly rules that the guard didn't cause the tsarina to survive.

3 Conjoined cases

Yet, the problem of isomorphs resurfaces as the problem of *conjoined cases*. Say, some normality theory solves an instance of the problem of isomorphs by assigning different normality evaluations to the values of some variables. Conjoin the two cases by these variables in a way that any witness that allows the theory to vindicate one intuition forces the theory to contradict the other intuition. You're produced a *conjoined case*—a counterexample to the theory. The literature lists at least four classic pairs of isomorphs: overdetermination and bogus prevention (explained above), causal and non-causal omissions, a cause and background condition, and early preemption and a short circuit. I'll use the pairs to produce four conjoined cases; one advantage of this strategy is that I'll reuse models of the conjoined cases that proponents of normality theories espouse themselves. I'll still work with *act-nrm*, but the argument works against any other normality theory that behaves similarly.

As the first conjoined case take *the tsarina*:

[3] The spy realizes that tea in the royal samovar is poisonous. She cannot let the empire descend into chaos. She has no time. She can replace only one cup, the tsar's or the tsarina's. She replaces the tsarina's. The tsarina's eager guard adds an antidote to her harmless tea. Meanwhile, the revolutionary finds his way to the tsar's chambers and poisons the already deadly tea. The tsar dies. The tsarina survives.

a. The revolutionary caused the tsar to die,

b. but the guard didn't cause the tsarina to survive.

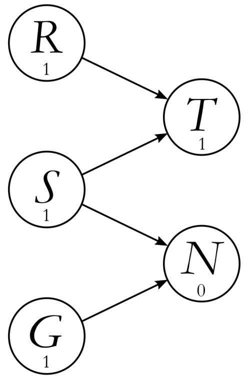

Figure 3. The tsarina. (Overdetermination and bogus prevention conjoined.)

As atomic events, take: $R = 1$ if the revolutionary poisons the tsar's cup, $R = 0$ if not; $T = 1$ if the tsar dies, $T = 0$ if not; S = -1 if the spy replaces the tsar's cup, $S = 1$ if the tsarina's; $S = 0$ if none; $N = 1$ if the tsarina dies, $T = 0$ if not; $G = 1$ if the guard adds the antidote to the tsarina's cup, $G = 0$ if not (fig. 3). Equations

$$R \leftarrow 1, \quad S \leftarrow 1, \quad G \leftarrow 1, \quad T \leftarrow S \geq 0 \vee R = 1, \quad N \leftarrow S \leq 0 \vee G = 0 \tag{4}$$

yield that $R = S = G = T = 1$ and $N = 0$.

The case conjoins overdetermination and bogus prevention by (what in the analyses of the original isomorphs serves as) the contingency variable. As analyzing overdetermination shows, to account for [1], $R = 1$ causing $T = 1$, you need a witness where the spy replaces the tsar's cup (S = -1), so that whether the tsar dies counterfactually depends on whether the revolutionary acts. Per *nrm*, this witness must be at least as normal as the actual solution, where the spy replaces the tsarina's cup (S = -1). But now on the standard solution to bogus prevention [2], $G = 1$ not causing $N = 0$, you must block any witness where the spy doesn't replace the tsarina's cup, including ones where she replaces the tsar's cup. Per *nrm*, no such witness can be at least as normal as the actual solution. The moral: however you assign normality to the values of G, you'll sacrifice one of the judgments.

For the second counterexample, recall the distinction between causally efficacious and inefficacious omissions (McGrath 2005). First, the original case and its standard solution from normality:

one neighbor but not the other agrees to water your plants while you're away. Neither waters the plants, but only the neighbor that agreed to help caused the plants to wilt. The standard solution from normality (Halpern & Hitchcock 2015:439) is to use a three-node model to represent the behavior of the neighbors and the fate of the plants. Assume an action is less normal than an omission unless the agent was expected to act (e.g., in virtue of a promise or a norm); if so, reverse the order. Therefore, the first neighbor's not watering the plants is less normal (as an instance of promise breaking) than not watering, but the second neighbor's not watering is more normal than watering. This way, you can use a witness where the first neighbor waters the plants but not one where the second does.

With this analysis in mind, consider *the garden*:

[4] Three neighbors share a garden. Melchior grows frankincense, Balthazar grows myrrh, and Caspar doesn't care. One day, Melchior and Balthazar are leaving for a trip. Melchior asks Caspar to turn on the sprinklers while he's gone; Balthazar isn't on speaking terms with Caspar. While the two travel, Caspar forgets to turn on the sprinklers. The plants wilt.

a. Caspar's failure to turn on the sprinklers caused the frankincense to wilt,

b. but it didn't cause the myrrh to wilt.

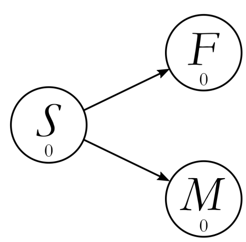

Figure 4. The garden. (A causal and non-causal omission conjoined.)

Take a natural model of the situation. Atomic events: $S = 1$ if Caspar turns on the sprinklers, $S = 0$ if not; $F = 1$ if the frankincense thrives, $F = 0$ if it wilts; $M = 1$ if the myrrh thrives, $M = 0$ if it wilts (fig. 4). Equations

$$S \leftarrow 0, \quad F \leftarrow S, \quad M \leftarrow S \tag{5}$$

yield that $S = F = M = 0$.

If you deem activating the sprinklers ($S = 1$) at least as normal as the corresponding omission ($S = 0$), you'll save [4a] but lose [4b] because the solution where Caspar turns on the sprinklers counts as at least as normal as the one where he doesn't and thus can serve as a witness for both judgments. If you deem activating the sprinklers less normal than (or incomparable with) the corresponding omission, you'll block any witnesses and thus save [4b] but lose [4a]. I've run this case by my peers. Although most agree with both [4a] and [4b], some (few) deny the later. If you side with them, that's fine; I have more cases.

The third counterexample exploits the distinction between causes and background conditions. Again, first consider the original case and its standard solution from normality: striking a match but not the presence of oxygen caused the flame even though the flame required oxygen. To account for the claim, take a three-node model where the variables represent whether the match is stricken, whether oxygen is present, and whether the match lights up. Take striking to be less normal than not striking but the presence of oxygen more normal than its absence (Halpern & Hitchcock, 2015:441). This way you have a witness for the claim that striking but not oxygen caused the flame.

With that in mind, consider *Rhodes*:

[5] Day by day, the Colossus' bronze skin turns green a little, and day by day, the Rhodians set the torch alight.

a. Air causes the bronze to patinate,

b. but it doesn't cause the torch to burn.

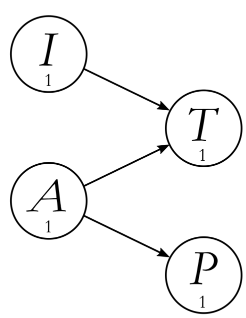

Figure 5. Rhodes. (A cause and background condition conjoined.)

For focus, assume that the model represents a particular day on Rhodes. Atomic events: $I = 1$ if the Rhodians attempt to light the torch, $I = 0$ if not; $A = 1$ if air is present, $A = 0$ if not; $T = 1$ if the torch burns, $T = 0$ if not; $P = 1$ for whether the bronze patinates, $P = 0$ if not (fig. 5). Equations

$$I \leftarrow 1, \quad A \leftarrow 1, \quad T \leftarrow I \wedge A, \quad P \leftarrow A \tag{6}$$

yield that $I = A = T = P = 1$.

To reuse Halpern and Hitchcock's strategy to account for [5b], you would need to treat the presence of air as more normal than its absence, which means that a solution where the air is absent ($A = 0$) can't serve as a witness. But that means you won't account for [5a], which requires a witness where air is absent.

The last case conjoins early preemption (Hitchcock 2001) with a short circuit (Hall 2007, Hitchcock 2007). First consider the original cases. *Early preemption*: the quisling poisons the tsar's soup, but if he hadn't, the spy would have. The quisling but not the spy caused the tsar's death. *Short circuit*: the spy laces the patriarch's soup with poison but only if the quisling adds an antidote first. The patriarch survives, but the quisling didn't cause it. The cases share the counterfactual structure: the spy's action depends on the quisling's, and both actions determine together what happens to the third character in the case. The standard solution from normality is this. In early preemption, treat the spy's (non-actual) poisoning the soup as less normal than her (actual) refraining from doing so. In the short circuit, treat the spy's (actual) poisoning the soup as at least as normal as her (non-actual) refraining from doing so. Normality assignments in both cases follow from the principle, espoused by Halpern and Hitchock (2015:451), that actual effects of actual causes—i.e., events on the output of an endogenous structural equation, when it's fed actual events—are always normal.

With this in mind, consider *the patriarch*:

[6] The tsar has ruled together with the patriarch. Not for long anymore. The the spy and the quisling are set to kill him. The tsar is allergic to tarragon and the patriarch to paprika. They die if they eat them. But if the patriarch eats paprika with tarragon, he will recover—tarragon will cure him.

The tsar and the patriarch feast together. Soup will be served. The quisling can spice it with tarragon or retreat. If the quisling goes through, the spy will add paprika. Otherwise, she will spice the soup with tarragon. Maybe to remove suspicion from the patriarch, who planned the murder? No one can tell.

The traitor adds paprika in the soup. The spy adds tarragon. This is the last feast for the tsar. But not for the patriarch—he lives on.

a. The quisling lacing the soup with tarragon caused the tsar to die,

b. but it didn't cause the patriarch to survive.

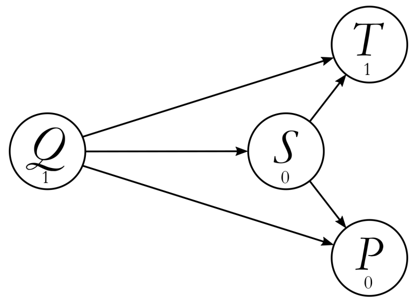

Figure 6. The patriarch. (Early prevention and a short circuit conjoined.)

Atomic events: $Q = 1$ for if the quisling poisons the soup with tarragon, $Q = 0$ if not; $S = 1$ if the spy adds tarragon; $S = 0$ if paprika; $T = 1$ if the tsar dies, $T = 0$ if he survives; $P = 1$ if the patriarch dies, $P = 0$ if he survives (fig. 6). Equations

$$Q \leftarrow 1, \quad S \leftarrow \neg Q, \quad T \leftarrow Q \vee S, \quad P \leftarrow \neg Q \wedge \neg S \tag{7}$$

yield that $Q = T = 1$ and $S = P = 0$.

To account for [6a], you need a witness where the quisling retreats ($Q = 0$) and the spy adds paprika to the soup ($S = 0$, the actual value) because only then does the tsar survive. But to account for [6b], you must block this very solution as a witness, lest you get the counterfactual dependence of the patriarch surviving on the quisling lacing the soup with tarragon.

Now, intuitions in these counterexamples would be saved if normality evaluations were effect-relative; e.g., if you deem Caspar's omission normal when entertained as a cause of the myrrh wilting but abnormal when entertained as a cause of the frankincense wilting, you'll save both [4a] and [4b]. But no conceptual analysis of normality leaves room for this kind of relativism (Bear & Knobe, 2017; Hitchcock & Knobe, 2009; Wysocki, 2020). Sure, normality is context dependent (snow is abnormal in Marib and normal in Reykjavik), but when taken as a part of a specific situation, an event is either normal or abnormal. To implement effect-relativity, for every variable, you would have to order its range not with one partial order (*at least as normal*) but with a partial order for every other variable from the domain (*at least as normal relative to X*). Moreover, some theories (Halpern & Pearl 2005, Halpern 2016) allow for non-atomic effect events. If so, such a modified theory would also need to tell whether the presence of oxygen is normal when entertained as a cause of the conjunctive event of bronze patinating *and* the torch burning ($P = 1 \wedge T = 1$). Not only is this solution unmanageably complex, but, more importantly, it doesn't seem to track any concept or a meaningful feature of the world.

4 Dealing with gerrymandering and other tricks

The cases invite objections. One: *I shouldn't conjoin the models in the first place*, as they are used for evaluating distinct causal judgments. The objection won't do. I'm not conjoining models, I'm conjoining situations, and then model them with a single model. Nothing wrong with that. On the standard procedure, you represent a situation with a model and then use it to derive any causal judgment that can be formulated with the model's variables. And indeed, a single model can and has been used to yield multiple causal judgments. For instance, in the standard treatment of the early preemption case (the QST part of the model from fig. 6), the same model is used to arrive at two judgments: that the quisling, but not the spy, caused the tsar to die.

The second objection: although it's fine to model each conjoined case with one model, *the models I use are inapt*. I could recover the judgments if I used more (or different) variables. When conjoining overdetermination and bogus prevention (fig. 3), for instance, I could use the extended model for bogus prevention (fig. 2) instead of the three-node original. The objection won't do. Notice that I'm conjoining models standardly used in the literature to represent this type of cases. If my model is inapt, then by analogy so is the original model for bogus prevention (fig. 1). That is, this objection capitulates to Blanchard and Schaffer's (2017) argument that invoking normality is unnecessary because once you account for all relevant events, isomorphs turn out not to be isomorphic after all.

Now—you may press—even though I build my cases out of other cases, this doesn't automatically mean that the best model of the conjoined cases conjoins the original models. Conjoining cases, that is, might not be so innocent, and might warrant adding new variables. This objection, I think, unduly shifts the burden. If there are better models of these situations that allow normality theories to yield the correct judgment, then it's on the proponent of the objection to produce such cases *and* to explain why I can't use the ones I use. The latter postulate is crucial. It's not enough only to claim that there are other apt models; one needs to show that I'm mistaken using mine.

Still, I'll say more. Although there are no uncontroversial sufficient aptness conditions for models, there are uncontroversial necessary conditions (Blanchard & Schaffer 2017, Hitchcock 2001, Halpern & Hitchcock 2010, Halpern & Pearl 2005, Pearl 2009, Spirtes et al. 2001). I'll just list them. The model entails only true counterfactuals. Atomic events denoted by values of different variables aren't conceptually related (e.g., one variable can't represent whether the tsar lives and another whether he's dead because a state where he lives while being dead is incoherent). Different values of the same variable denote inconsistent events (the quisling can't poison and not poison the soup at the same time). Variable values shouldn't represent events that we consider too remote and thus aren't willing to take seriously (e.g., that the spy bilocates and replaces both poisoned cups at the same time). No combination of values of different variables may represent an incoherent state of affairs (Ross & Woodward ms.).[6] The models that represent the conjoined cases clearly satisfy these criteria.

Consider one more: you may add a variable to a model if it doesn't change what Halpern and Hitchcock (2010:395) call the topology of the model, by which they mean at least the number of directed paths to the effect variable and whether the paths cross and where. Since can view the model of any conjoined case as the standard model of one of the isomorphs with an extra variable that doesn't change the topology of the original sub-model, my models satisfy the requirement. E.g., my model for the Rhodes case (fig. 5) consists of the model for the standard cause-background condition case, IAT, and an extra variable, P, that doesn't alter any of the original relations between the values of I, A, and T. As the extra variable doesn't alter the structure of the sub-model, if you think that the standard model of the isomorph is apt, you shouldn't doubt that my models are apt as well.

In general, it seems plausible that the threshold for forbidding models should be rather high: if a model satisfies all these necessary criteria, you need a strong reason to reject a model.[7] That a theory yields an incorrect causal judgment when applied to a model doesn't count as a strong reason.

The third objection: fair enough, I've shown that the standard treatment of isomorphs doesn't work on my puzzles. But *there are other theories*. For instance, Gallow (2021) proposes a theory that you can calibrate to your needs. I won't present his entire theory, which is rather complex; what's relevant for my cases is that you can demand that the contrast cause be more normal than the actual cause, and/or that the *contrast effect* (i.e., the value the effect variable assumes under the contrast cause and contrast contingency) is more normal than the actual effect. This theory handles, e.g., the garden and the tsarina case. Assume that the actual effect must be less normal than its contrast, and consider the garden (fig. 4). Declare the frankincense wilting ($F = 0$) more normal than flourishing ($F = 1$), but the myrrh wilting ($M = 0$) less normal than flourishing ($M = 1$), and you'll save both [4a] and [4b]. You can handle the tsarina analogously (fig. 3). Declare the tsar's and tsarina's deaths

[6]The most convincing reason I see for this requirement an intervention cannot bring about a state of affairs that cannot exist, and interventions can bring about any combination of values of different variables.

[7]I'm not alone in thinking this; see (Gallow ms.:10). In fact, I disagree (and so does Gallow 2021) with Halpern and Hitchcock's topology consideration: if a model satisfies all other criteria listed here, I can't see why we should care how it relates to other apt models of the situation. Whether a model is apt should only be a matter of the relation between this model and the situation modeled. Still, I list the topology criterion just in case someone finds it crucial.

less normal than them surviving, and you'll save both [3a] and [3b].

The first solution is arbitrary, for F and M are symmetrical; the second goes against Halpern and Hitchcock's desideratum that outcomes of structural equations are prima facie normal. But it would be nice to have a stronger argument. And there is one.

Consider the *symmetric garden* (fig. 7), a variation on the garden case (fig. 4).

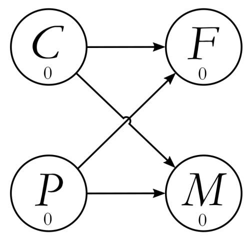

Figure 7. The symmetric garden.

We all misremember. There was no Caspar but two unreliable neighbors: Cas promised Melchior to water his frankincense; Par promised Balthazar to water his myrrh. There still was one sprinkler system, though the neighbors could activate it independently from their houses–if at least one neighbor had activated the system, it would water the garden at night. Let C and P stand for whether Cas and Par activate the system, and F and M for whether the frankincense and myrrh flourish or wilt (fig. 7). Both Cas and Par forgot to turn on the sprinkler:

$$C \leftarrow 0, \quad P \leftarrow 0, \quad F \leftarrow C \vee P, \quad M \leftarrow C \vee P. \tag{8}$$

Intuitions to save: Cas caused frankincense but not myrrh to wilt; Par caused myrrh but not frankincense to wilt. If $F = 0$ is at least as normal as $F = 1$, a theory that requires actual effects be less normal than their contrasts won't accommodate the first claim; if $F = 0$ is less normal than $F = 1$, the theory won't accommodate the second claim. A theory that requires actual causes to be less normal than their contrasts will have the same problem when you try to assign normality to C's and P's values.[8]

Consider next *smuta* (fig. 8), a variation on the tsarina case (fig. 3).

[8]In this and the original garden case, you may think that a better model would be one that distinguishes between activating the sprinklers and the sprinklers watering the garden. In the current case, that would mean adding a variable, S, for whether the sprinklers are on. The equations would state: $C \leftarrow 0$, $P \leftarrow 0$, $S \leftarrow C \vee P$, $F \leftarrow S$, and $M \leftarrow S$. Maybe this is a more natural model; fortunately for my argument, the model proves as problematic.

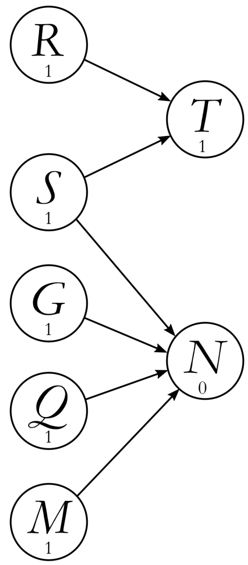

Figure 8. Smuta. (Overdetermination, bogus prevention, and efficacious prevention.)

The new setup is like the original but for two additional actors: the quisling and the maid. The quisling attempts to kill the tsarina ($Q = 1$) by poisoning the wires of her earrings, but her maid always wipes the wires ($M = 1$) before the tsarina puts the earrings on. That is,

$$\begin{gathered} R \leftarrow 1, \quad S \leftarrow 1, \quad G \leftarrow 1, \quad T \leftarrow S \geq 0 \vee R = 1, \\ Q \leftarrow 1, \quad M \leftarrow 1, \quad N \leftarrow (S \leq 0 \vee G = 0) \vee (Q = 1 \wedge M = 0)\,. \end{gathered} \tag{9}$$

The case is easier to entertain once you notice that the QNM triangle constitutes a case of efficacious prevention: the maid prevented the tsarina's death, which is to say that the maid caused the tsarina to survive. If the tsarina's surviving ($N = 0$) is at least as normal as her dying ($N = 1$), a theory that requires actual effects to be less normal than their contrast loses the judgment that the guard didn't cause her to survive. If her surviving is less normal than her dying, the theory loses the judgment that the maid did cause the tsarina to survive.

Now say you're instead working with a theory on which actual causes must be less normal than their contrasts. To save the judgment that the guard didn't cause the tsarina to survive, you can take his adding the antidote to the tea as more normal his not doing it, which means that the latter won't serve as a contrast. But you can modify the original case (fig. 3) accordingly. Say, for instance, that the guard's supervisor will be impressed if the guard adds an antidote to the cup (I won't specify the model, as it is obvious). And indeed, he does, hence causing the supervisor to be impressed. If $G = 1$ is at least as normal than $G = 0$, you'll mishandle the latter claim; if less normal, you'll mishandle the claim that the guard didn't cause the tsarina to survive.

The fourth objection: abandon any pretense to being principled and just notice that *you can always order the witnesses to preserve the judgments*, although this ordering won't be a function of the normality of the component events. Let's engage in some gerrymandering, that is. For illustration, again revisit the tsarina (fig. 3). Let any witness where $S = -1$, $R = 0$, and $G = 1$ be as normal as the actual assignment, and the claim that $R = 1$ causes $T = 1$ is vindicated. Let any witness where $S \leq 0$ and $G = 0$ be less normal than the actual assignment, which means that there's no witness for $G = 1$ to be a cause of $N = 0$. This order allows *nrm* to account for both target causal claims.

The objection won't do. The proposal is *ad hoc*: there's no principled reason to treat an assignment where $S = -1$, $R = 0$, and $G = 1$ as at least as normal than the actual assignment where $S = 1$, $R = 1$, $G = 1$, and this one as more normal than any assignment where $S \leq 0$, $R = 1$, and $G = 0$. No causal theory that employs normality allows you to order assignments arbitrarily. The order on assignments is always a function of the order on variable values *unless*—should we listen to Halpern and Hitchcock (2013, 2015:435, Halpern 2016:80-1)—you have some special justification, such as some events frequently occurring together. Moreover, Halpern and Hitchcock allow for this caveat only if the normality of events doesn't yield an unequivocal ordering over the target assignments, i.e., if the assignments would be otherwise incomparable. None of that is available to the objection in question, which lets me dismiss it with a light heart.[9] In fact, Halpern and Hitchcock (2010:402) warn against any such maneuver: "[including normality] raises the worry, however, that this gives the modeler too much flexibility. After all, the modeler can now render any claim that A is an actual cause of B false, simply by choosing a normality order that assigns the actual world [i.e., assignment] a lower rank than any world [i.e., witness] needed to satisfy [the definition of causation]." Again, their remedy is to derive the order from facts about frequency, moral and social norms, and norms of proper functioning.

The fifth objection: *old news, old news*; all I did was producing more examples, but no one claims that normality can solve all problems (Halpern 2016:90). Just add my cases to the pile to be solved by another improvement to the counterfactual theory. I disagree. The main problem is that my conjoined cases, as I've stressed already, reuse cases that normality was supposed to handle. Say, that a future iteration of the counterfactual theory accounts for the conjoined cases. When applied

[9]Once you allow for arbitrary orderings of assignments, you can do away with equations and counterfactual dependence, for you'll be able to handle any case whatsoever. No one who espouses a normality theory holds that view. Hüber (2013) has a theory where he unites normality and counterfactual dependence into one measure, but his orderings are far from arbitrary.

to the conjoined cases, this future theory won't appeal to normality (at least in the same way that current normality theories do). But if so, then you should be able to apply the theory to original isomorophs without appealing to normality either, for the models of isomorphs are just sub-models of conjoined cases.

Now, this argument isn't bulletproof. It might turn out that the future theory handles, say, the claim that the quisling caused the tsar's death (T), [6a], by taking into account what happens to the patriarch (P). And if so, then this solution wouldn't apply to the original isomorophs. Although this would be a weird theory, for T and P don't causally influence each other, I can't rule such a theory out.[10] Still, I feel like I've offered enough justification to bring up conjoined cases as a family of cases that are especially problematic for normality theories in ways that, say, voting scenarios (Livengood 2013) or trumping preemption (Schaffer 2000) isn't, even though normality theories typically have problems with handling these other cases too.

The argument is new in another respect too. While I agree with Blanchard and Schaffer (2017) that incorporating normality as a measure over variable values doesn't work, I'm not so convinced by their positive proposals. In their treatment of bogus prevention (fig. 2), they add a variable that represents "whether or not any neutralization occurs" (Blanchard & Schaffer 2017:201). Adding the variable seems to violate one of the criteria, for the event of 'neutralization occurring' seems to require, as a matter of conceptual truth, the neturaliz*er* and the neutraliz*ed*. That is, an assignment where there's no poison ($S = 1$) but neutralization occurs ($C = 1$) seems incoherent. If the tea isn't poisoned, there's no meaningful way in which the antidote can counteract the nonexistent poison.[11] Therefore, unlike Blanchard and Schaffer, I don't maintain that the standard models of isomorophs are inapt.

Moreover, recall the original case of causally efficacious and inefficacious omissions: neither neighbor watered your plant, but only the one who agreed to water it caused it to wilt. Blanchard and Schaffer propose that we shouldn't let the problematic variable into the model in the first place. I.e., per the criteria of good modeling, we shouldn't take seriously the possibility that the neighbor who you didn't ask to water your plant could break into your house and do it. While I agree with this treatment of the original case, an analogous move won't work for the garden (fig. 4) and especially

[10] Rather: I can't rule out this theory *here*. Take a causal claim $C = c$ causes $E = e$. I think it is a plausible constraint on a structural-equations theory of causation that accounting for this causal claim shouldn't appeal to variables that neither are on the chosen path from C to E nor parent an on-path variable. This constraint would forbid appealing to P while showing why $Q = 1$ causes $T = 1$ in the patriarch case. However, I do not wish my current argument to require adopting this constraint.

[11] Ross and Woodward (ms.) and Gallow (ms.) probably would concur.

the symmetric garden case (fig. 7). In my models of these cases, all values in the variables' ranges represent serious possibilities. Therefore, despite what Blanchard and Schaffer would probably have hoped for, it seems that the structural-equations framework alone cannot handle these cases.

5 Conclusion

Altogether, I have offered seven counterexamples to current normality theories. These aren't just any counterexamples—the theories were designed, I daresay, to handle this sort of cases. That's not all. I have offered a recipe for brewing more problems. Say, some theory handles a particular pair of cases by assigning different normality evaluations to the values of some variables.[12] Conjoin the two cases by these variables in a way that any witness that allows the theory to vindicate one intuition forces the theory to contradict the other intuition. You've produced a counterexample. I don't claim this will always work, but so far the recipe seems successful.

Now, I'm not saying that normality doesn't play a role in lay causal judgments. The empirical evidence for that is too strong (see, e.g., Alicke et al. 2015, Clarke 2015, Gerstenberg & Icard 2019, Henne et al. 2017, Henne et al. 2021, Hitchcock & Knobe 2009, Knobe & Fraser 2008, Kominsky & Phillips 2019). Nor am I saying *here* that normality shouldn't play a role in causal judgments. I am saying, though, that the way current normality theories incorporate normality doesn't work. Assigning normality to values of variables overlooks the fact that the same event can be causally efficacious toward one effect and in inefficacious toward another (non)effect. Branding events as normal/abnormal or default/deviant simply doesn't leave room for this obvious possibility.

6 References

1. Alicke, M. D., Mandel, D. R., Hilton, D. J., Gerstenberg, T., Lagnado, D. A. (2015). Causal conceptions in social explanation and moral evaluation: A historical tour. *Perspectives on Psychological Science* 10(6):790–812.

2. Bear, A., Knobe, J. (2017). Normality: Part descriptive, part prescriptive. *Cognition* 167:25–37.

3. Beckers, S., Vennekens, J. 2018. "A principled approach to defining actual causation.", *Synthese* 195, 835–862.

[12] In fact, the cases may but needn't be each other isomorphs.

4. Beebee, H., Hitchcock, Ch., Price, H. 2017. *Making a Difference*, Oxford.

5. Blanchard, T., Schaffer K. (2017). Cause without Default. In: Beebee, Hitchcock, Price, 2017, 175-214.

6. Clarke, R., Shepherd, J., Stigall, J., Waller, R., Zarpentine, C. (2015). Causation, norms, and omissions: A study of causal judgments. *Philosophical psychology* 28(2):279-293.

7. Gallow, D. (2021). A Model-Invariant Theory of Causation. *Philosophical Review* 130(1):45-96.

8. Gallow, D. (ms.). Dependence, Defaults, and Needs.

9. Gerstenberg, T., Icard, T. (2019). Expectations affect physical causation judgments. *Journal of Experimental Psychology: General* 149(3):599–607.

10. Glymour, C., Wimberly, F. (2007). Actual causes and thought experiments. In: Campbell, O'Rourke, Silverstein, *Causation and Explanation*, Cambridge, 43-67.

11. Hall, N. (2007). Structural equations and causation. *Philosophical Studies* 132(1):109–136.

12. Halpern, J. 2016. *Actual Causality*. Cambridge.

13. Halpern, J., Hitchcock, C. (2010). Actual causation and the art of modeling. In: Dechter, Geffner, Halpern (eds.), *Causality, Probability, and Heuristics: A Tribute to Judea Pearl*, London, 383-406.

14. Halpern, J., Hitchcock, C. (2013). Compact Representations of Extended Causal Models. *Cognitive Science* 37(6):986-1010.

15. Halpern, J., Hitchcock, C. (2015). Graded Causation and Defaults. *British Journal for the Philosophy of Science* 66(2):413-457.

16. Halpern, J., Pearl, J. (2005). Causes and Explanations: A Structural-Model Approach, Part I: Causes. *British Journal for the Philosophy of Science*, 56(4), 843–87.

17. Henne, P., O'Neill, K., Bello, P., Khemlani, S., De Brigard, F. (2021). Norms Affect Prospective Causal Judgments. *Cognitive Science* 45:e12931.

18. Henne, P., Pinillos, Á., De Brigard, F. (2017). Cause by omission and norm: Not watering plants. *Australasian Journal of Philosophy* 95(2):270–283.

19. Hiddleston, E. (2005). A causal theory of counterfactuals. *Noûs*, 39:632–657.

20. Hitchcock, C. (2001). The Intransitivity of Causation Revealed in Equations and Graphs. *Journal of Philosophy*, 98(6), 273-299.

21. Hitchcock, C., Knobe, J. (2009). Cause and norm. *Journal of Philosophy* 106:587–612.

22. Hüber, F. (2013). Structural equations and beyond. *Review of Symbolic Logic* 6(4):709-732.

23. Knobe, J., Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In: Sinnott-Armstrong, *Moral Psychology* vol. 2, 441–448.

24. Kominsky, J., Phillips, J. (2019). Immoral professors and malfunctioning tools: Counterfactual relevance accounts explain the effect of norm violations on causal selection. *Cognitive Science* 43, e12792.

25. Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs* 13:455–476.

26. Livengood, J. (2013). Actual Causation in Simple Voting Scenarios. *Noûs* 47(2):316–345.

27. Maudlin, T. (2004). Causation, Counterfactuals, and the Third Factor. In: Collins, Hall, Paul (eds.), *Causation and Counterfactuals*. Cambridge.

28. McGrath, S. (2005). Causation by Omission: A Dilemma. *Philosophical Studies*, 123, 125–148.

29. Menzies, P. (2004). Causal models, token causation, and processes. *Philosophy of Science* 71(5):820-832.

30. Menzies, P. (2017). The Problem of Counterfactual Isomorphs. In: Beebee, Hitchcock, Price, 153–74.

31. Morris, A., Phillips, J., Gerstenberg, T., Cushman, F. (2019). Quantitative causal selection patterns in token causation. *PLoS One* 14(8), e0219704.

32. Paul, L., & Hall, N. (2013). *Causation: A User's Guide*. Oxford.

33. Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*, second edition, Cambridge.

34. Ross, L., Woodward, J. (ms.) Irreversible (One-hit) and Reversible (Sustaining) Causation. http://philsci-archive.pitt.edu/19836/1/Irr_Rev_Ross_Woodward.pdf

35. Schaffer, J. (2000). Trumping preemption. *Journal of Philosophy* XCVII(4):165-181.

36. Spirtes, P., Glymour, C., Scheines, R. (2001). *Causation, Prediction and Search, 2nd revised edition*, Cambridge.

37. Thomson, J. (2003). Causation: Omissions. *Philosophy and Phenomenological Research* 66(1):81–103.

38. Weslake, B. (2015). A Partial Theory of Actual Causation. *British Journal for the Philosophy of Science*.

39. Wolff, J. E. (2016). Using Defaults to Understand Token Causation. *The Journal of Philosophy* 113(1):5–26.

40. Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*, New York.

41. Wysocki, T. (2020). Normality: a Two-Faced Concept. *Review of Philosophy and Psychology* 11:689–716.