Commentary

Factual Difference-Making is Equivalent to a Counterfactual Theory

Forthcoming in Australasian Philosophical Review

Lennart B. Ackermans

31st May 2025

Abstract

In *Factual Difference-Making*, Holger Andreas and Mario Günther propose a theory of model-relative actual causation which performs remarkably well on a number of known problematic cases. They take this to show that we should abandon our counterfactual way of thinking about causation in favour of their factual alternative. I cast doubt on this argument by offering two similar theories. First, I show that the theory of Factual Difference-Making is equivalent to a partly counterfactual theory. Second, I give a fully counterfactual theory that makes the same judgments in the scenarios discussed by Andreas and Günther.

1.

In *Factual Difference-Making*, Holger Andreas and Mario Günther propose a theory of model-relative actual causation whose causal judgments satisfy the causal intuitions of many philosophers remarkably well. Unlike most existing theories, Andreas and Günther claim that their theory does not rely on counterfactual reasoning. Instead, it is based on the notion of *factual difference-making*, reminiscent of sufficiency conditions in the regularity approach to causation.

It is certainly a significant achievement of *Factual Difference-Making* that it gives intuitively plausible judgments for a large amount of varying scenarios. But the authors take this achievement to show that we should adopt a paradigm shift in the way in which we think of causation. As they state in the conclusion: "perhaps we shouldn't think of causes as counterfactual difference-makers. [...] Perhaps we should think of causes as factual difference-makers" (Andreas & Günther, forthcoming, p. 46).

In this commentary, I cast doubt on this conclusion by introducing a counterfactual theory of causation that is equivalent to *Factual Difference-Making*.

This equivalent counterfactual theory still uses conditions that are somewhat incongruous with the counterfactual approach. However, I also provide a closely related theory which is "fully counterfactual" and makes the same judgments in all cases discussed by Andreas and Günther.

Section 2 introduces notation and defines the two counterfactual theories. Section 3 discusses an example. Section 4 concludes. The appendix A provides proofs.

2.

I will adopt the notation and terminology from Andreas and Günther (forthcoming), with a few additions and clarifications. See their article for more elaborate definitions.

A causal model $\langle M, V\rangle$ consists of a set of structural equations M and a valuation V. For an equation $A = \varphi \in M$, A is a propositional variable and φ a propositional formula of other variables. The set of equations M is acyclic.[1] A valuation V is a set of literals stating which variables are true and false. For each $A \in V$ or $\neg A \in V$, A is a variable of M. A variable A of M is called *exogenous* when there is no equation $A = \varphi \in M$.

We say $\langle M, V\rangle \models \varphi$ if and only if the variable assignment of V to M is consistent and satisfies φ as a matter of classical logic. (See Andreas and Günther, forthcoming

[1] M is acyclic when there is no sequence of variables $A_1, \dots, A_n = A_1$ such that the for each equation $A_{i+i} = \varphi \in M$, the variable A_i appears in φ.

for a preciser definition.) When we have neither $\langle M, V\rangle \models \varphi$ nor $\langle M, V\rangle \models \neg\varphi$, we say that $\langle M, V\rangle$ is unsettled on φ. For a set of literals I such that for every $X \in I$, $\neg X \notin V$, the factual intervention on I is defined by

$$\langle M, V\rangle[I] = \langle M_I, V \cup I\rangle, \text{ where}$$
$$M_I = \{(A = \varphi) \in M \mid A \notin I \wedge \neg A \notin I\}.$$

The preceding definitions match those in Andreas and Günther (forthcoming). In addition, I will introduce the notion of a propagating intervention common in counterfactual theories.

Definition 1. Let $\langle M, V\rangle$ be a consistent causal model. For any set of literals I for variables of M, the *propagating intervention* on I, denoted $\langle M, V\rangle[\![I]\!]$, is the model obtained by intervening on the literals in I and propagating the changes to their descendants. More formally, let $Z \subseteq V \setminus I$ be the set of literals of nondescendants of I excluding I. We define

$$\langle M, V\rangle[\![I]\!] = \langle M_I, V_I\rangle, \text{ where}$$
$$M_I = \{(A = \varphi) \in M \mid A \notin I \wedge \neg A \notin I\},$$
$$V_I = \{\psi \mid \langle M_I, Z \cup I\rangle \models \psi \text{ and } (\psi = A \vee \psi = \neg A, \text{ for } A \text{ a variable of } M)\}.$$

I will say that V is a *full valuation* for a model $\langle M, V\rangle$ if V contains a literal for every variable in M.

The definition of causation by Andreas and Günther (*Factual Difference-Making*) is repeated in the appendix as definition 5. The following theory defines when an event C causes an event E in a partly counterfactual sense, denoted $\gg_3$. This theory is equivalent to Factual Difference-Making (proposition A.4).

Definition 2 ($C \gg_3 E$)**.** Let $\langle M, V\rangle$ be a consistent causal model. $\langle M, V\rangle \models C \gg_3 E$ if and only if there is full valuation V' and $M' \subseteq M$ such that

(0) M' does not contain a structural equation for C.

(1) $\neg C, \neg E \in V'$ and $\langle M', V'\rangle$ is consistent,

(2) there is no V'' such that $V' \cap V \subset V'' \cap V$ and $\neg C, \neg E \in V''$ and $\langle M', V''\rangle$ is consistent.

(3) $\langle M', \varnothing\rangle[V' \cap V][C] \models E$,

(4) the structural equation of each descendant of C is in M', and

(5) for any literal $C' \in V' \setminus V$ whose variable is neither a descendant nor an ancestor of C, $\neg C'$ is more deviant than C'.

I will refer to condition (n) of $\gg_3$ as $\gg_3(n)$.

There are two ways in which the above theory might be said not to be a theory in the counterfactual tradition. First, condition $\gg_3(1)$ requires that there is a consistent model with $\neg C, \neg E \in V'$, whereas typical theories use the notion of a propagating intervention. However, since M' of condition (1) does not contain an equation for C, condition (1) is equivalent to the following interventionist condition:

(1') $\langle M', V\rangle [\![\neg C \cup I]\!] \models \neg E$ for some intervention set I of exogenous variables.

Using (1') would require a reformulation of conditions (2) and (3). I use (1) instead to stay close to the theory of Factual Difference-Making.

Second, condition $\gg_3(3)$ is similar to condition 3 of Factual Difference-Making (definition 5) and uses a model that is not fully specified, i.e., $\langle M', \varnothing\rangle$. Thought experiments in which variables have no value – rather than a true or false value – are unheard of in counterfactual theories and typical of the approach of Factual Difference-Making. We could, in theory, avoid this by iterating over consistent models with full valuations. That is, $\gg_3(3)$ is equivalent to the following:

(3') For all full valuations W such that $C \in W$, and $\langle M', W\rangle [V' \cap V]$ is consistent, we have $\langle M', W\rangle [V' \cap V] \models E$.

However, this is just a rephrasing, and still not entirely consonant with the counterfactual approach. Alternatively, we can replace condition (3) of $\gg_3$ with a condition that uses the notion of a propagating intervention. This yields the following theory that is similar to Factual Difference-Making, though not quite equivalent.

Definition 3 ($C \gg_4 E$)**.** Let $\langle M, V\rangle$ be a consistent causal model. $\langle M, V\rangle \models C \gg_4 E$ if and only if there is full valuation V' and $M' \subseteq M$ such that

(0,1,2,4,5) As in definition 2.

(3) $\langle M', V'\rangle [V' \cap V][\![C]\!] \models E$,

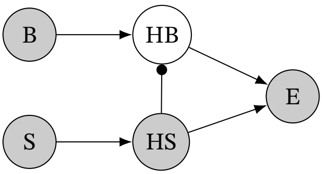
The above definition gives a theory of actual causation that is fully counterfactual (and I think this is an improvement). It implies Factual Difference-Making, that is, if A causes B in the sense of $\gg_4$, then A causes B in the sense of Factual Difference-Making (proposition A.3). Since condition $\gg_4(3)$ is weaker, there are cases in which A causes B according to it but not according to Factual Difference-Making (proposition A.5).

However, $\gg_4$ and Factual Difference-Making make the same judgments about causes and non-causes in all cases discussed by Andreas and Günther (forthcoming), that is, early preemption, late preemption, the boulder scenario, bogus prevention, and omissions. The achievements of Factual-Difference-Making thus carry over to this counterfactual theory.

All cases I could find in which $\gg_4$ and Factual-Difference-Making disagree involve complex equations with at least three input variables and a combination of AND and OR (see proposition A.5). (My intuitions as to which theory is correct are unclear for such cases.) This leads me to suspect that $\gg_4$ and Factual-Difference-Making are equivalent when we restrict our attention to causal models that only contain "simple" equations that do not combine AND an OR, and have negations applied only directly to variables (and not to expressions of multiple variables). However, I was unable to prove this.

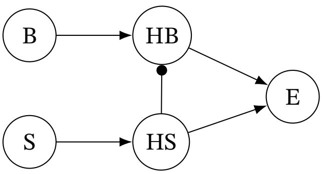
3.

As an example, consider a model of late preemption. Billy and Suzy throw rocks at a bottle. Suzy's rock hits the bottle first, shattering it. If Suzy's rock had not hit the bottle, then Billy's rock would have hit it and shattered it. The variables B and S denote Billy and Suzy respectively throwing their rock. The variable HS and HB denote Suzy's and Billy's rock hitting the bottle. The variable E denotes the shattering of the bottle. The model $\langle M, V \rangle$ is as follows.



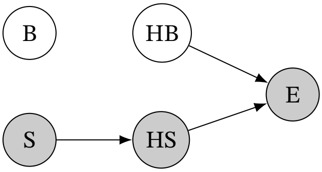
M:	V:
$HB = B \wedge \neg HS$	$B, \neg HB, S,$
$HS = S$	HS, E
$E = HB \vee HS$	

To see that Suzy's throw (S) is a cause of the bottle shattering (E), consider the following counterfactual model $\langle M', V' \rangle$.



M':	V':
$HB = B \wedge \neg HS$	$\neg B, \neg HB,$
$HS = S$	$\neg S, \neg HS,$
$E = HB \vee HS$	$\neg E$

This model clearly satisfies condition $\gg_3(0)$ and $\gg_3(4)$. It satisfies condition $\gg_3(1)$ and $\gg_3(2)$, since the only consistent model with fewer changes of actual variables is the model that that has B, which would satisfy E. Condition $\gg_3(3)$ requires that we remove the equation for HB in the above model, set S to true, HB to false, and unsettle all other variables. The resulting model satisfies E, so we have $\langle M, V\rangle \models S \gg_3 E$. For $\gg_4(3)$ we similarly remove the equation for HB and do a propagating intervention with $I = \{S\}$. This gives the model $\langle M', V'\rangle[\{\neg HB\}][\![C]\!]$ depicted below, which satisfies E. Hence, we also have $\langle M, V\rangle \models S \gg_4 E$.



$HS = S$ $\qquad$ $\neg B, \neg HB,$

$E = HB \vee HS$ $\qquad$ S, HS, E

To see that Billy's throw (B) is not a cause of the bottle shattering (E) according to both counterfactual theories, consider that the model $\langle M', V'\rangle$ above is also the only consistent model satisfying $\neg B$ and $\neg E$. After removing the equation for HB, setting B to true changes nothing (on both a factual and propagating intervention), so B does not cause E according to both theories.

4.

Andreas and Günther argue that we need a paradigm shift in our thinking of actual causation towards considering causes as factual difference-makers. I hope to have undermined their main argument for this position by showing that the successes of Factual Difference-Making can be enjoyed by counterfactual theories. In closing, I offer two additional reasons to be sceptical that the authors' proposed shift in thinking is required.

First, how we should think of actual causation depends at least in part on the metaphysics of actual causation, a matter not settled by *Factual Difference-Making*. That is, it is desirable that the way in which we think of causes coincides with what causes actually are. A theory of model-relative causation can be said to be metaphysically accurate only if the metaphysical interpretation of the model is correct. It is clear from other writings (e.g., Andreas and Günther, 2024) that the authors favour a regularity-theoretic interpretation of causal models. But one can just as well use the theory of Factual Difference-Making while interpreting the model's equations in a counterfactual sense.

Moreover, the specifics of the theory of Factual Difference-Making itself (disregarding model interpretation) do not seem to matter metaphysically, that is, they do not point to or require a particular metaphysical interpretation. We can interpret the model-relative conditions of Factual Difference-Making in the sense favoured by the authors, but as I've shown, we can equally give the theory's conditions a counterfactual formulation. (And perhaps, we need not interpret the theory's conditions at all.)

Hence, the success of Factual Difference-Making does not speak in favour or against any metaphysical theory of causation. This means that an important consideration when deciding how we should think of causation has not been addressed by the authors (nor by my commentary).

Second, another desideratum of the way in which we think of causes is that doing so is sufficiently easy. As humans, we are used to counterfactual reasoning and can practice it with little effort. If we take the author's proposal seriously, we should develop a new sort of reasoning faculty. When faced with a causal scenario (Billy and Suzy throw a rock at a bottle) we should imagine first that the relevant events neither occur nor not occur. We should then ask ourselves whether the proposed effect (the bottle shatters) is sure to occur if the proposed cause occurs (Billy throws), holding fixed certain events. This way of thinking – imagining events that are neither occurring nor absent – is surely very difficult, and it would be better if we can avoid that outcome.

A.

The proofs will regularly make use of the following notion of a submodel.

Definition 4. A causal model $\langle M, V\rangle$ is called a *submodel* of a causal model $\langle M', V'\rangle$ if and only if $M \subseteq M'$ and $V \subseteq V'$. Then the latter is called a *supermodel* of the former.

A useful property of submodels, regularly used below, is that consequences of the submodel are consequences of the supermodel. That is, given that $\langle M, V\rangle$ is a submodel of $\langle M', V'\rangle$ and both are consistent, $\langle M, V\rangle \models \varphi$ implies $\langle M', V'\rangle \models \varphi$.

Below I repeat the definition of causation as factual difference-making from Andreas and Günther (forthcoming).

Definition 5 ($C \gg_1 E$)**.** Let $\langle M, V\rangle$ be a causal model. $\langle M, V\rangle \models C \gg_1 E$ if and only if there is $V' \subseteq V$ and $M' \subseteq M$ such that

(1) $\langle M', V'\rangle$ is unsettled on C and E,

(2) there is no $V'' \subset V$ such that $V' \subset V''$ and $\langle M', V''\rangle$ is unsettled on C and E,

(3) $\langle M', \varnothing\rangle[V'][C] \models E$,

(4) the structural equation of each descendant of C is in M', and

(5) for any literal $C' \in V \setminus V'$ whose variable is neither a descendant nor an ancestor of C, C' is more deviant than $\neg C'$.

This definition can be simplified by only considering models that have the equation for C removed. (This will make the subsequent proofs easier.)

Definition 6 ($C \gg_2 E$)**.** Let $\langle M, V\rangle$ be a causal model. Then $\langle M, V\rangle \models C \gg_2 E$ if and only if there is $V' \subseteq V$ and $M' \subseteq M$ such that

(0) M' does not contain a structural equation for C.

(1)–(5) As in definition 5.

Proposition A.1. *Let $\langle M, V\rangle$ be a causal model. Then $C \gg_1 E$ if and only if $C \gg_2 E$.*

Proof. If $\langle M, V\rangle \models C \gg_2 E$, then it is trivial that $\langle M, V\rangle \models C \gg_1 E$. So suppose that $\langle M, V\rangle \models C \gg_1 E$.

Let M', V' be such that $\gg_1$(1)-(5) of definition 5 are satisfied. Let M'' be the model obtained by removing the equation for C from M'. V' is unsettled on C and E with respect to M'' since it is unsettled with respect to M' (which contains up to one extra equation). So let $V'' = V' \cup W$ be a valuation for M'' that is *minimally* unsettled on C and E. We have $V \setminus V'' \subseteq V \setminus V'$, so condition $\gg_2(5)$ of definition 6 is satisfied with respect to V''. We now have that M'' and V'' satisfy conditions $\gg_2(0)$, $\gg_2(1)$, $\gg_2(2)$, $\gg_2(4)$, and $\gg_2(5)$. It remains to be shown that condition $\gg_2(3)$ is satisfied.

We have $V' \subseteq V''$ and both $\langle M'', \varnothing\rangle[V''][C]$ and $\langle M', \varnothing\rangle[V'][C]$ have the equation for C removed. Hence, $\langle M'', \varnothing\rangle[V''][C]$ is a supermodel of $\langle M', \varnothing\rangle[V'][C]$. The latter model satisfies E, and so the former must as well. $\square$

Proposition A.2. *Let $\langle M, V\rangle$ be a causal model such that $\langle M, V\rangle \models C \wedge E$. If we have $\langle M, V\rangle \models C \gg_2 E$, then we have $\langle M, V\rangle \models C \gg_3 E$ and $\langle M, V\rangle \models C \gg_4 E$.*

Proof. Suppose $\langle M, V\rangle \models C \gg_2 E$. Let M', V_1' such that $\gg_2(0)$-(5) of definition 6 are satisfied. By condition $\gg_2(3)$, we have $\langle M', \varnothing\rangle[V_1'][C] \models E$. This is a submodel of $\langle M', V_1'\rangle[C]$, so the latter model satisfies E or is inconsistent. It can't be inconsistent, since it is a submodel of $\langle M, V\rangle$, which is consistent. Since M' does not contain an equation for C, this model equals $\langle M', V_1' \cup \{C\}\rangle$, so we have $\langle M', V_1' \cup \{C\}\rangle \models E$.

Suppose that $\langle M', V_1' \cup \{\neg C\}\rangle \models E$. Then combined with the above we have $\langle M', V_1'\rangle \models E$, contradicting $\gg_2(1)$. Hence, $\langle M', V_1' \cup \{\neg C\}\rangle$ is unsettled on E or satisfies $\neg E$. Hence, there is a full valuation $V_2' = V_1' \cup W$ such that $\neg C, \neg E \in W$ and such that $V_2' \cap V$ is maximal in the sense of $\gg_3(2)$. Hence, M' and V_2' satisfy $\gg_3(1)$, $\gg_4(1)$, $\gg_3(2)$, and $\gg_4(2)$ (definition 2 and 3).

Suppose $W \cap V$ is non-empty. The model $\langle M', V_1' \cup (W \cap V)\rangle$ is a submodel of both $\langle M', V\rangle$ (which satisfies C and E) and a submodel of $\langle M', V_1' \cup W\rangle$ (which satisfies $\neg C$ and $\neg E$). So $\langle M', V_1' \cup (W \cap V)\rangle$ is unsettled on C and E, contradicting $\gg_2(2)$. Hence, we have $W \cap V = \varnothing$.

Hence, we have $V_2' \cap V = V_1'$. Condition $\gg_3(3)$ follows directly from $\gg_2(3)$.

Let Z consist of the exogenous literals (including C) of $\langle M', V_2'\rangle[V \cap V_2'][\![C]\!]$. Then the models $\langle M', V_2'\rangle[V \cap V_2'][\![C]\!]$ and $\langle M', Z\rangle[V \cap V_2']$ are "equal", i.e., they contain the same equations and satisfy the same literals.[2] Moreover, we have $V_1' = V \cap V_2'$, so the model $\langle M', \varnothing\rangle[V_1'][C]$ is a submodel of $\langle M', Z\rangle[V \cap V_2']$. Since the former model satisfies E, the latter model satisfies E. Condition $\gg_4(3)$ follows.

If $C' \in V_2' \setminus V$, then $\neg C' \notin V_1'$ and $C' \notin V$, so $\neg C' \in V \setminus V_1'$. So by $\gg_2(5)$, $\neg C'$ is more deviant than C'. Hence, V_2' satisfies $\gg_3(5)$ and $\gg_4(5)$.

Since M' satisfies $\gg_2(0)$ and $\gg_2(4)$, it satisfies $\gg_3(0)$, $\gg_4(0)$, $\gg_3(4)$, and $\gg_4(4)$. $\square$

[2] This is because a causal model's true literals are fully determined by a valuation of all exogenous variables.

Proposition A.3. *Let $\langle M, V\rangle$ be a causal model such that $\langle M, V\rangle \models C \wedge E$. If we have $\langle M, V\rangle \models C \gg_3 E$, then we have $\langle M, V\rangle \models C \gg_2 E$.*

Proof. Let M', V_3' such that conditions $\gg_3$(0)-(5) of definition 2 are satisfied.

Let $V_1' = V \cap V_3'$. The model $\langle M', V_1'\rangle$ is a submodel of both $\langle M', V\rangle$ and $\langle M', V_3'\rangle$. The former satisfies C and E while the latter satisfies $\neg C$ and $\neg E$. Hence, $\langle M', V_1'\rangle$ is unsettled on C and E. Therefore, M' and V_1' satisfy condition $\gg_2$(1).

Let $X \in V \setminus V_1'$ and let Z be any partial valuation such that $W = V_1' \cup Z \cup \{X\}$ is a full and consistent valuation of M'. Then $\langle M, W\rangle$ cannot satisfy both $\neg C$ and $\neg E$, since then W would violate $\gg_3$(2). Moreover, $\langle M, W\rangle$ cannot satisfy both C and $\neg E$, since then $\langle M, \varnothing\rangle[V_1'][C]$ would be a submodel of $\langle M, W\rangle$, but the former satisfies E by $\gg_3$(3). Hence, every complete valuation containing $V_1' \cup \{X\}$ satisfies E. It follows that V_1' is minimally unsettled, satisfying $\gg_2$(2).

From $\gg_3$(3) it follows immediately that $\gg_2$(3).

Let $C' \in V \setminus V_1'$ be a literal that is neither a descendant nor ancestor of C. Then $C' \in V$ and $C' \notin V_3' \cap V$, so $\neg C' \in V_3' \setminus V$. By $\gg_3$(5), C' is more deviant than $\neg C'$. Hence, V_1' satisfies $\gg_2$(5).

By $\gg_3$(0) and $\gg_3$(4), M' satisfies $\gg_2$(0) and $\gg_2$(4). □

By proposition A.1, A.2, and A.3 we have equivalence of $\gg_3$ and $\gg_1$.

Proposition A.4. *Let $\langle M, V\rangle$ be a causal model such that $\langle M, V\rangle \models C \wedge E$. Then we have $\langle M, V\rangle \models C \gg_3 E$ if and only if $\langle M, V\rangle \models C \gg_1 E$.* □

Proposition A.5. *There exists a causal model $\langle M, V\rangle$ such that we have $\langle M, V\rangle \models C \wedge E$, $\langle M, V\rangle \models C \gg_4 E$, but $\langle M, V\rangle \not\models C \gg_1 E$.*

Proof. Let $M = \{E = CAB \vee C\neg A\neg B \vee \neg C(A \vee B)\}$ and let $V = \{A, B, C, E\}$. The only consistent counterfactual model with $\neg C$ and $\neg E$ is $V' = \{\neg A, \neg B, \neg C, \neg E\}$. The model following a propagating intervention on C satisfies E, so we have $\langle M, V\rangle \models C \gg_4 E$.

The two minimally unsettled models on C and E are $V_1' = \{A\}$ and $V_2' = \{B\}$. There are consistent valuations $\{C, A, \neg B, \neg E\}$ and $\{C, \neg A, B, \neg E\}$. Hence, adding C does not satisfy E in either case. We have $\langle M, V\rangle \not\models C \gg_1 E$. □

References

Andreas, H., & Günther, M. (2024). A lewisian regularity theory. *Philosophical Studies*, *181*(9), 2145–2176. https://doi.org/10.1007/s11098-024-02149-z

Andreas, H., & Günther, M. (forthcoming). Factual difference-making. *Australasian Philosophical Review*.