Philosophy of Science (2026), **00**, 1–21
doi:10.1017/psa.2026.10216

PSA PHILOSOPHY OF SCIENCE ASSOCIATION
PHILOSOPHY OF SCIENCE

ARTICLE

The Case For Time in Causal DAGs

Alexander Gilbert Reisach[1], Alberto Suárez[2], Sebastian Weichwald[3] and Antoine Chambaz[1]

[1]Université Paris Cité, France, [2]Universidad Autónoma de Madrid, Spain and [3]University of Copenhagen, Denmark
Corresponding author: Alexander Gilbert Reisach; Email: alexander.reisach@math.cnrs.fr

(Received 03 July 2025; revised 20 April 2026; accepted 20 April 2026)

Abstract

We make the case for incorporating a notion of time into causal directed acyclic graphs (DAGs). We demonstrate that nontemporal causal DAGs are ambiguous and obstruct justification of the acyclicity assumption. Assuming that causes precede effects, causal relationships are relative to the time order, and causal DAGs require temporal qualification. We propose a formalization via *composite* causal variables that refer to quantities at one or multiple time points. We emphasize that the acyclicity assumption requires different justifications depending on whether the time order allows cycles. We conclude by discussing implications for the interpretation and applicability of DAGs as causal models.

1. Nontemporal causal DAGs

Causal directed acyclic graphs (DAGs), as popularized by Pearl (2000, 2009), have been conceived as nontemporal causal models. Pearl (2009, sec. 7.5.1) states, on the matter of using time to determine causation:

> The reliance on temporal information has its price, as it excludes a priori the analysis of cases in which the temporal order is not well-defined.

Time is real, however, and there is no escaping it and whatever price it really may exact, whether temporal information is relied upon or not. We therefore argue that it is best to use time explicitly, and set out to explore the limitations and opportunities it reveals. In doing so, we rely on the principle that *causes precede effects*. This principle is fundamental to the definition of causality. In proposing the method of path coefficients, the precursor to graphical causal models, Wright (1934) states that "only such paths are tried which are appropriate in direction in time." The Bradford Hill criteria for causality (Hill 1965), widely used in epidemiology, refer to the same principle as "temporality." Granger (1969), Suppes (1970), and Shoham (1990) rely on the time order between cause and effect for their theories of causality. Throughout their foundational work on causal discovery, Spirtes et al. (1993); Spirtes et al. (2001) emphasize the existence of a time order. Similarly, the canonical work on graphical

© The Author(s), 2026. Published by Cambridge University Press on behalf of Philosophy of Science Association.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

causal models (Pearl 2009, e.g. sec. 7.5.1), as well as (Peters et al. 2017, e.g. ch. 10), concur that causality follows the time order.

Despite its central role in causality, the notion of time is absent from the formal treatment of causal DAGs, which define variables and edges without an explicit reference to time. In contrast to the potential outcomes framework of statistical causality (Splawa-Neyman 1932; Rubin 1974), causal DAGs have been developed explicitly for "nontemporal statistical data" (Pearl 2009, ch. 2). This expression covers two distinct properties. First, the distributions compatible with a DAG do not depend on time. Second, and perhaps more controversially, a DAG imposes no constraints on the temporal structure within realizations of such a distribution beyond those entailed by the causal order (see Section 3 for details). This offers a powerful unified theory regardless of whether the time order prevents cycles, or whether they are possible but absent (Pearl 2009, sec. 7.5.1). However, the general nature of this approach obscures the connection between time and causation, as well as the rationale for the acyclicity assumption. As a result, causal DAGs have been used widely, including in contexts such as neuroimaging (Ramsey et al. 2017) or gene regulatory networks (Lee et al. 2019), where acyclicity is questionable. In many cases, the search for nontemporal acyclic causal relationships has led to variables without physically well-defined interventions, such as raising cities or changing their climate in the popular *altitude* → *temperature* example (e.g. Peters et al. 2017, sec. 2.1). The absence of time has also prompted the use of time proxy variables like *age* or *hour of day* (e.g. Mooij et al. 2016), and even the modelling of time itself as a cause (Huang et al. 2015), despite the obvious issue of defining corresponding interventions.

These problems are symptomatic of the tension that arises from the contrast between the essential role of time in causation and the nontemporality of causal DAGs: It is unclear whether a given nontemporal DAG relates quantities at all, any, or some specific time points, and hence how acyclicity should be justified. Moreover, many real-world variables of interest such as *economic inflation*, *precipitation*, or *clinical depression* are complex temporal aggregations. For these reasons, we make the case for incorporating the notion of time into the definition of variables in causal DAGs in an explicit and principled fashion.

1.1. Nontemporal causal DAGs are ambiguous

In causal DAGs, the values taken by the random variables refer to quantities of a type of object in some external system modelled by the DAG. In this sense, we say as a shorthand that a random variable *refers to* or *measures* quantities. We refer to a full assignment of values to all random variables of the model as a *realization*. Consider the DAG $A \to H$ between a random variable A measuring the intake of the drug aspirin and a random variable H measuring the symptom headache, a popular example in the literature (see, e.g., Shpitser and Pearl 2008; Imbens and Rubin 2015; Pearl and Mackenzie 2018). The present edge encodes that (i) *A causes H*, and the absent edge in the opposite direction encodes that (ii) *H does not cause A*. We rely on the interventionist view of causation (Woodward 2005) typically used in connection with causal DAGs (Pearl 1995, sec. 4; 2009). Since causes precede effects, the interventionist reading of (i) entails the statement

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press



(a) Aspirin before headache. **(b)** Headache before aspirin.

Figure 1. Timeline of instantiated aspirin (•) and headache (•) variables.

(I) ***first*** *intervening on A may change the* ***subsequent*** *distribution of H*

for some unspecified time distance. However, (ii) does not necessarily entail that

(II) ***first*** *intervening on H does not change the* ***subsequent*** *distribution of A.*

The ambiguity arises because it is unclear whether nontemporal DAGs make any statement about time orders that do not follow the causal order. To illustrate this point, consider the two different timelines of instantiated (i.e. time-point-specific) aspirin and headache variables shown in figure 1. It follows from the causal order that the DAG models a time order of aspirin before headache, as shown in figure 1(a), but it is unclear whether it also models the inverse order of headache before aspirin, as shown in figure 1(b). If the DAG is taken to describe only aspirin before headache, then the absence of an effect of H on A follows purely from the time order, and the DAG provides no information about statement (II). If it is taken to describe both time orders, then the absent edge does entail statement (II) for some unspecified time distance. In light of the ambiguity, it may seem tempting to interpret the causal order in $A \to H$ as an implicit instantiation of the form (A_{t_0}, H_{t_1}) with $t_0 < t_1$. However, such an interpretation would remain ambiguous for variables not connected by a directed causal path. Thus, since causal effects depend on time order and may also depend on time distance, the interpretation of nontemporal DAGs is inherently ambiguous.

1.2. Justifying the acyclicity assumption requires a notion of time

In a nontemporal DAG, it is unclear which time point or set of time points is involved in the definition of each of the random variables. This issue is particularly problematic when it comes to evaluating the validity of the acyclicity assumption. We summarize the relationship between time order and acyclicity in Remark 1, which is a direct consequence of the principle that causes precede effects within realizations.

Remark 1. A causal cycle between two variables requires that at least one of them refers to quantities at multiple time points.

For illustration, consider a version of the running aspirin and headache example with multiple possible measurements for each, as shown in figure 2. Assume that aspirin causes subsequent headache and vice versa (aspirin relieves headaches, so the presence of a headache would be a reason to take aspirin). If A and H each refer to a quantity at a single time point, the DAG between them is acyclic by the time order.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

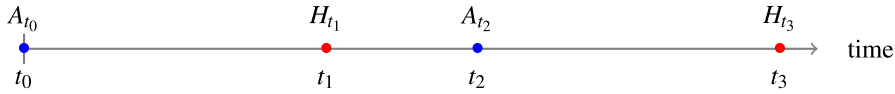


Figure 2. A timeline with multiple possible measurement times for aspirin and headache.

A cycle is possible if one or both variables refer to quantities at multiple time points. We call such variables *composite* variables. Composite variables can refer to quantities at randomly drawn time points, and they can aggregate quantities at multiple, possibly random, time points. For instance, A may be uniformly drawn from $\{A_{t_0}, A_{t_2}\}$, or be the sum $A_{t_0} + A_{t_2}$. More generally, the example highlights that a DAG cannot relate the generic concepts of aspirin and headache, but only temporal instantiations thereof. For example, the DAG may change from $A \to H$ for $A := A_{t_0}$ and $H := H_{t_1}$, to $H \to A$ for $H := H_{t_1}$ and $A := A_{t_2}$, to the cyclic causal graph $A \rightleftarrows H$ if the variables are composed of multiple time points. It is thus unclear what causal statements nontemporal DAGs actually imply, and whether acyclicity follows from the time order or is an additional assumption. To address this issue, we propose a formalism to specify the temporality of variables in causal DAGs.

1.3. Contribution

We discuss related work on time in causal modelling and causal time-series models in Section 2. We propose a new formalization of time for causal variables of varying temporal complexity in Section 3, and characterize two separate conditions for acyclicity: acyclicity by a time order that prevents cycles, and acyclicity by the absence of causal cycles where the time order allows them. In Section 4 we illustrate our formalism in different temporal configurations, and demonstrate how it can be used to reason systematically about causal cycles, appropriate time scales, and potential faithfulness violations. In Section 5.1 we discuss the implications for causal discovery, effect estimation, and the applicability of causal DAGs. Section 6 summarizes our conclusions, and Appendix A provides additional detailed formalizations.

2. Related work

2.1. Causal DAGs

DAGs are the central tool for representing causal relationships using graphical models (Pearl 2009). A DAG $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of vertices $\mathcal{V}$, and a set of directed edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ such that if $(X, Y) \in \mathcal{E}$, then there is a directed edge from X to Y. We call a *directed path* an ordered collection of directed edges such that the end point of the kth edge is the start point of the $(k+1)$th edge. By definition, DAGs are acyclic, that is, there are no directed paths that begin and end in the same vertex. Each vertex represents a random variable, and their joint distribution is assumed to be Markov with respect to the DAG (e.g. Pearl 2009, definition 1.2.2; Peters et al. 2017, definition 6.21). A causal DAG represents an external causal system (Dawid 2010, sec. 5), and a directed edge $(X, Y) \in \mathcal{E}$, also denoted as $X \to Y$, indicates that X is a direct cause of Y within the variables of the graph. Hence, a DAG encodes *causal structure*, and a DAG with a set of additional parameters is a *causal model* (Pearl 2009, definitions 2.2.1 and 2.2.2). The superclass of *structural causal models* (Pearl 2009, definition 7.1.1) does not

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

in principle require acyclic structures, but is nonetheless most often used in conjunction with DAGs (see, e.g., Pearl 2009, sec. 1.2; Peters et al. 2017, sec. 6.2), which are the central object of study and interest in the field of graphical causal models due to their favorable properties.

2.2. Critical perspectives on causal DAGs

Conventionally, causal DAGs relate random variables, which are *type* variables, whereas the causal order operates within the process by which these variables are realized, that is, at the *token* level. For details on the type vs. token debate, see (Williamson 2004, sec. 2.3) and Spohn (2012, ch. 14; 2019). The common justification of assuming the existence of a time order for the acyclicity of a causal DAG implicitly elevates a token-level property to the type level. Our work provides the motivation and formalism for aligning the temporal restrictions that act on the token level with the definition of type-level variables in a principled fashion.

Our contribution ties into a series of critical perspectives on causal DAGs. Causal DAGs allow for far-reaching causal insights because they are a restrictive model class. In his methodological critique, Dawid (2010) points out that the strong assumptions underlying the model class require adequate justification. Greenland (2010) argues that causal DAGs rely on null hypotheses that are implausible in many applications. Aalen et al. (2016) question whether DAGs between measurements taken at discrete time points are useful for understanding underlying continuous causal processes. Notably, the works by Greenland (2010) and Aalen et al. (2016) do not question the directionality of edges or the acyclicity assumption. Writing from the perspective of the applied sciences, they naturally assume that the causal variables refer to quantities measured at certain time points, giving rise to a time order. This contrasts with the idea of nontemporal *equilibrium* variables predominant in the causal DAG literature (see, e.g., Pearl 2009, ch. 2; Spirtes et al. 2010, sec. 5; Peters et al. 2017, remark 6.7). Spohn (2010, sec. 3.6) notes the problem that nontemporal variables can refer to items of varying temporal positions. Motivated by the same observation, Weinberger (2019) proposes a temporal interpretation for equilibrium causal DAGs, and Weinberger (2023) discusses the temporality of interventions. We adopt the applied sciences' pragmatic view that measurements are necessarily time specific, and take Spohn's critique to its logical conclusion: the causal relationships encoded in DAGs depend on the temporal nature of the variables. This motivates our development of a formalism that can encode a wide range of possible temporal configurations, highlights the need to justify acyclicity, and clarifies the interpretation of causal DAGs.

2.3. Temporal graphical causal models

There are, of course, many approaches in the literature on graphical causal models that already incorporate a notion of time. The principal difference of our contribution is that we do not propose a temporal approach as a special or alternative case, but rather argue that all causal DAGs are temporal in nature if one accepts that causes precede effects.

In the existing literature, cyclic graphical causal models (Bongers et al. 2021) offer a way to essentially preserve the nontemporal approach while acknowledging that

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

variables that aggregate quantities at multiple time points may be characterized by cyclic causal relationships. Stern (2021) points out that considering the temporal order may greatly simplify causal discovery. A rich literature (e.g., Peters et al. 2013; Malinsky and Spirtes 2018; Runge 2018; Assaad 2021) applies dynamic Bayesian networks (Dean and Kanazawa 1989; Murphy 2002) as causal models for the classical time-series setting involving regular measurements of a dynamic process. Similarly, Hansen and Sokol (2014) show that stochastic differential equations can be seen as a continuous-time generalization of time-series DAGs. Our approach constitutes a bridge between nontemporal and time-series causal DAGs. We highlight that, even if realizations are mutually independent and no matter the level of aggregation, causal variables always have a time dimension which is crucial for interpreting causal relationships and for assessing the validity of the acyclicity assumption. Conceptually, the perspective we propose is closest to that of the potential outcomes framework (Splawa-Neyman 1932; Rubin 1974), and aligns with the definition of causal variables in Imbens and Rubin (2015, sec. 1.2) and Hernán and Robins (2020, sec. 22.4). Our contribution extends the existing literature by considering variables that can refer to quantities at multiple and varying time points. In doing so, we make use of and provide a temporal perspective on the idea behind *cluster* and *group* variables as studied in Anand et al. (2023) and Wahl et al. (2024).

3. Time in causal DAGs

We set out to formalize a notion of time that is sufficiently flexible to capture variables of different temporal complexity. To see that there is no contradiction between temporal variables and time-invariant causal mechanisms, consider that a realization of the variables in a causal DAG in the sense of Pearl (2009) results from applying deterministic causal mechanisms to a draw of exogenous noise variables (following the conception of Laplace 1814). Each variable in a DAG is typically equipped with its own independent noise source, so the resulting model is still indeterministic. The causal mechanisms are invariant to the starting time of their application, yet their causal effects unfold in time within each realization. We refer to the timeline relative to the starting point of one realization as *within realization*, and distinguish it from the *between-realization* timeline relating the starting points of different realizations. For there to be a causal effect between two variables in distribution, their within-realization time order must follow their causal order with positive probability.

To illustrate the distinction between the within-realization and between-realization timelines, figure 3 provides an example of four realizations of two variables X and Y. Each realization contains one value per variable that refers to a quantity at a time point marked by $\times$. This corresponds to a setting where the variables X and Y can refer to one of multiple instantiated variables (that is, variables that refer to a quantity at a specific time point) with positive probability.

Note that the causal order restricts the within-realization time order (in a probabilistic fashion) regardless of whether realizations are dependent or independent from one another. This adds an important nuance to the common view that independent and identically distributed realizations imply the absence of time structure (see, e.g., Peters et al. 2013, sec. 1; 2017, remark 6.7). In light of figure 3, the

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

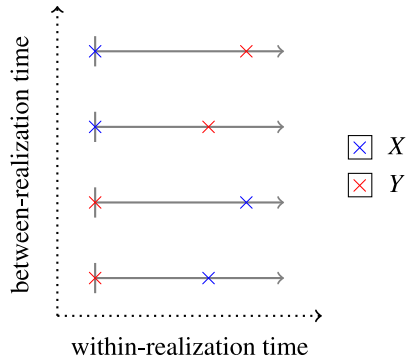


Figure 3. Four realizations with their respective timelines.

problems of nontemporal DAGs outlined in Section 1.1 and 1.2 can be understood as resulting from the lack of specifying the within-realization time points that variables correspond to. The notion of time we aim to formalize in the remainder of this section is therefore within-realization.

3.1. A formal notion of time for causal variables

Beginning with variables that refer to a quantity at a single time point, we develop our formalism to allow for composite variables, which can refer to quantities at multiple, possibly random, time points.

3.1.1. Single time point

In the simplest case, each variable refers to a quantity at exactly one time point relative to a realization's timeline. By default, we take the time point t of a variable $\mathbb{X}_t$ to be the time point at which it is measured, but other specifications are equally possible (e.g., proxy measurements of past quantities). For simplicity, each within-realization timeline starts at 0. This formalization captures the notion of time employed in the literature on causality in time series. In this simple model, the acyclicity assumption holds by the order of time. To probe the acyclicity assumption where it may be violated and to model richer settings, we must allow for variables that can refer to quantities at multiple time points.

3.1.2. Multiple time points

In all but the most strictly controlled settings, variables may refer to multiple and varying relative time points in different realizations. In observational medical data, for example, the time structure between measurements may vary from patient to patient depending on the availability of medical personnel and equipment. To formalize such settings in the form of composite variables, we introduce additional objects. We index them with the same lower-case letter if they are used in the definition of the same composite variable. Let $\mathcal{T}_x \subset \mathbb{R}_+$ be a countable set of possible time points. We assume that for each time point $t \in \mathcal{T}_x$ there exists a corresponding real-valued random variable $\mathbb{X}_t$, which we capture collectively by the real-valued stochastic process $(\mathbb{X}_t)_{t\in\mathcal{T}_x}$. For a fixed integer $k_x \in \mathbb{N}^*$, let $\mathbb{T}_x \subseteq \mathcal{T}_x$ be a

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Table 1. The objects used for formalizing a composite variable $X_{\mathbb{T}_x}$ with an explicit notion of time

Function	Object	Interpretation
Single time point	$\mathbb{X}_t$	Single variable at time point t
Multiple time points	k_x	Number of time points per realization of $X_{\mathbb{T}_x}$
	$\mathscr{T}_x$	Set of possible time points of which $\mathbb{T}_x$ is a random subset
	$\mathbb{T}_x$	Actual time points; random subset of k_x time points of $\mathscr{T}_x$
Aggregating time points	f_x	Aggregation function of $(\mathbb{X}_t)_{t\in\mathbb{T}_x}$

random variable taking values in the set of k_x-element subsets of $\mathscr{T}_x$. The process $(\mathbb{X}_t)_{t\in\mathbb{T}_x}$ can then be used to formalize a random selection from multiple different time points. In each realization of the stochastic process, the set of time points is drawn according to the law of $\mathbb{T}_x$. This law can be thought of as a law over measurement procedures.

3.1.3. Aggregating multiple time points

Many variables of practical interest are aggregations of quantities at multiple time points, e.g. in the form of averages over time or differences between time points. For a prominent example, consider the historical debate regarding the effect of smoking on lung cancer (see, e.g., Spirtes et al. 2001, sec. 9.5). In this debate, *smoking* refers to a composite variable that describes a behavioral pattern over a time period. To formalize such settings, let $f_x\colon \mathbb{R}^{k_x} \to \mathbb{R}$ be a fixed real-valued aggregation function such that $f_x(\mathbf{x}) \neq f_x(\mathbf{x}')$ if $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^{k_x}$ differ in any single component. This ensures that an intervention on one temporal component also results in a change in the aggregate variable. In the smoking–cancer example, a real-valued smoking variable might, for instance, be the difference in total cigarettes smoked between two time points divided by the length of the time period.

3.1.4. Putting it all together

On the basis of the objects introduced and summarized in Table 1, we can now formalize a composite variable $X_{\mathbb{T}_x}$ that can refer to a quantity at a single time point, or to a random selection or aggregation of quantities at multiple time points. For any valid k_x and any distribution of $\mathbb{T}_x$, we define the composite causal variable

$$X_{\mathbb{T}_x} := f_x((\mathbb{X}_t)_{t\in\mathbb{T}_x}), \tag{1}$$

which explicitly references corresponding time points in contrast to its nontemporal counterpart X. Owing to the essential role of time in causality and the problems due to its absence in conventional causal DAGs outlined in Section 1.1 and 1.2, we choose to keep the time index explicit even after the aggregation into a single real-valued random variable. The objects in Table 1 encompass a wide range of possible variable definitions, though in principle our formalism allows for further generalization, for example by making k_x a random variable. Moving forward, all causal variables will be composite variables of the form given in equation (1).

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

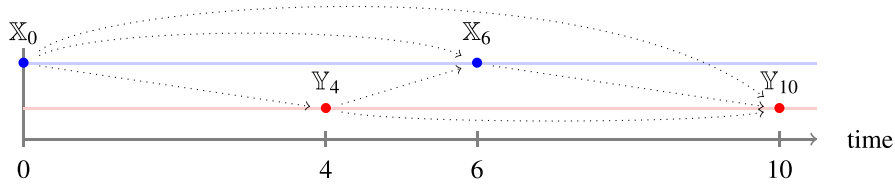


Figure 4. Atomic causal DAG.

3.1.5. The atomic causal DAG

Following the time series DAG literature (e.g. Hyttinen et al. 2012; Bongers et al. 2021), we assume the existence of a causal DAG between variables that refer to quantities at single time points. We refer to this DAG as the *atomic* causal DAG, and within it all causes precede their effects (see definition A.1). From the atomic causal DAG it is possible to derive a directed graph between composite variables, which may or may not be acyclic. Interventions on $X_{\mathbb{T}_x}$ may target $\mathbb{X}_t$ at any time point t that can appear with positive probability in $\mathbb{T}_x$. We say that $X_{\mathbb{T}_x}$ causes another causal variable $Y_{\mathbb{T}_y}$ if there exists (t_x, t_y) that occurs jointly with positive probability and $\mathbb{X}_{t_x}$ causes $\mathbb{Y}_{t_y}$ in the atomic DAG, without considering self-loops (see definition A.2; for a detailed treatment of inference rules over cluster variables we refer to Anand et al. 2023). We denote the existence of a causal path between two variables by $\rightsquigarrow$. If the causal path consists of a single edge, then $\rightsquigarrow$ can be replaced by a straight arrow $\rightarrow$.

3.1.6. Running example

To illustrate how time-point-specific variables relate to composite causal variables, consider a generic medical treatment–symptom setting following the formalization introduced above with a treatment process $(\mathbb{X}_t)_{t\in\mathcal{T}_x}$ with $\mathcal{T}_x = \{0, 6\}$, and a symptom process $(\mathbb{Y}_t)_{t\in\mathcal{T}_y}$ with $\mathcal{T}_y = \{4, 10\}$. The timeline is shown in figure 4. For visual clarity, the temporal components of each composite variable are drawn on separate lines belonging to the same within-realization timeline, as indicated by the vertical connection. The dotted arrows show causal relationships that are compatible with the time order (shown only in this figure).

On the basis of the stochastic processes, one can define composite variables $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$. For instance, one may define the composite treatment $X_{\mathbb{T}_x} := (\mathbb{X}_0 + \mathbb{X}_6)/2$, and the composite outcome $Y_{\mathbb{T}_y}$ as the identity evaluated either at $\mathbb{Y}_4$ or $\mathbb{Y}_{10}$ with equal probability. Table 2 summarizes this in terms of the objects in Table 1. For more examples, see Section 4.

3.2. A tale of two acyclicities

The temporality of the variables in a causal DAG plays a central role in the justification of the acyclicity assumption. We partition the acyclicity assumption into two separate and individually sufficient conditions with distinct characteristics: acyclicity by a time order that prevents cycles, and acyclicity by the absence of causal cycles where the time order allows them. The latter is a stronger assumption that requires additional justification. A detailed formalization of the concepts introduced here can be found in Appendix A. We begin by defining *time-acyclicity* as acyclicity by

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Table 2. Composite variable definitions

	$\mathscr{T}$ (possible)	k	$\mathbb{T}$ (actual)	f
$X_{\mathbb{T}_X}$	$\{0, 6\}$	2	$\{0, 6\}$	average
$Y_{\mathbb{T}_Y}$	$\{4, 10\}$	1	$\{4\}$ or $\{10\}$	identity

the time order. It is the reason for acyclicity in many common examples in the literature (see, e.g., Spirtes et al. 2001; Pearl 2009; Spohn 2012, ch. 14).

Time-acyclicity (see definition A.4 for a formal version)

Time-acyclicity holds between two causal variables if and only if the latest time point of one always precedes the earliest time point of the other.

Discussions of the acyclicity assumption typically focus on the presence or absence of causal cycles such as feedback loops (see, e.g., Spirtes et al. 2000; Brouillard et al. 2025). Causal cycles can only occur if there is no time-acyclicity, which leads us to the definition of *effect-acyclicity* as acyclicity by the absence of possible causal relationships that would lead to cycles.

Effect-acyclicity (see definition A.5 for a formal version)

Effect-acyclicity holds between two causal variables if and only if time-acyclicity does not hold for them, yet there is still no causal cycle between them.

A special case of effect-acyclicity between two causal variables is given if there is no way for one to affect the other, regardless of the temporal relationship between them. We refer to this condition as *total effect-acyclicity* and define it as follows.

Total effect-acyclicity (see definition A.6 for a formal version)

Total effect-acyclicity holds between two causal variables if and only if there is no causal cycle between them regardless of which of their possible time points are selected.

Total effect-acyclicity is the type of acyclicity with the most far-reaching consequences. For a nonempty graph, total effect-acyclicity describes a generic asymmetry irrespective of the definition of a particular composite variable. Such asymmetries are a key reason for the appeal of causal models in other fields, in particular in machine learning (Schölkopf 2022). Unfortunately, they may be hard to justify since they are statements about the entire underlying stochastic processes (compare the formal definitions in Appendix A).

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

3.2.1. The contrast to acyclicity for nontemporal variables

In our formalism, the correspondence of causal variables to time points is explicit. Therefore, an acyclic causal relationship between two variables such as $X_{\mathbb{T}_x} \rightsquigarrow Y_{\mathbb{T}_y}$ does not in general preclude that some $\mathbb{Y}_t$ may cause some $\mathbb{X}_{t'}$. Note the contrast to the paradigm of nontemporal DAGs, in which a path $X \rightsquigarrow Y$ categorically excludes any path $Y \rightsquigarrow X$, but it remains unclear if this is meant to refer to time- or effect-acyclicity for a particular sets of time points, or to a fundamental asymmetry in the sense of total effect-acyclicity.

4. Illustrative examples

We begin by discussing a series of examples illustrating our formalism and how to justify acyclicity using the concepts presented in Section 3.2. We use our running example from figure 4 consisting of treatment process $(\mathbb{X}_t)_{t\in\mathscr{T}}$ with $\mathscr{T}_x = \{0, 6\}$ and symptom process $(\mathbb{Y}_t)_{t\in\mathscr{T}_y}$ with $\mathscr{T}_y = \{4, 10\}$. Then, we show how our formalism can be used to reason systematically about causal cycles, time scales, and unobserved confounding. Finally, we demonstrate how it may help uncover faithfulness violations.

Example 1 (Single time point). For our first and simplest scenario, we define our composite variables to be a quantity at a single one of all possible time points. For instance, let $X_{\mathbb{T}_x} := \mathbb{X}_0$ and $Y_{\mathbb{T}_y} := \mathbb{Y}_{10}$, with $\mathbb{T}_x = \{0\}$, $\mathbb{T}_y = \{10\}$, $k_x = k_y = 1$, and both of f_x and f_y defined as the identity function of the variable in their singleton input set. This is an example of the time-point-specific variable definition also used in the literature on causality in time series. The example exhibits time-acyclicity, as is easy to see since the time order allows $X_{\mathbb{T}_x}$ to be a cause of $Y_{\mathbb{T}_y}$, but not vice versa.

Example 2 (Multiple time points). A more complex scenario emerges for variables that can refer to a quantity at a randomly chosen time point. For instance, let $X_{\mathbb{T}_x}$ be equal to either $\mathbb{X}_0$ or $\mathbb{X}_6$, and $Y_{\mathbb{T}_y}$ be equal to either $\mathbb{Y}_4$ or $\mathbb{Y}_{10}$. This is the case if $\mathbb{T}_x$ is either $\{0\}$ or $\{6\}$ with positive probability, and $\mathbb{T}_y$ is either $\{4\}$ or $\{10\}$ with positive probability, $k_x = k_y = 1$, and f_x and f_y are the identity as before. Here, time-acyclicity is no longer as easy to justify since, by the time order, $X_{\mathbb{T}_x}$ could cause $Y_{\mathbb{T}_y}$ and vice versa, e.g. if $\mathbb{X}_0 \rightsquigarrow \mathbb{Y}_4$ and $\mathbb{Y}_4 \rightsquigarrow \mathbb{X}_6$. Still, time-acyclicity may hold, e.g. if the joint event of $(\mathbb{T}_y = \{4\}, \mathbb{T}_x = \{6\})$ has probability 0, meaning that the symptom measurement always follows the treatment. If all combinations of time points have positive probability, acyclicity requires justification via effect-acyclicity, e.g. by the absence of a causal effect of $\mathbb{Y}_4$ on $\mathbb{X}_6$ in the atomic DAG.

Example 3 (Aggregation). Variables of interest may be aggregates of quantities at multiple time points. For instance, let $X_{\mathbb{T}_x} = (\mathbb{X}_0 + \mathbb{X}_6)/2$ and $Y_{\mathbb{T}_y} = (\mathbb{Y}_4 + \mathbb{Y}_{10})/2$, with $\mathbb{T}_x = \{0, 6\}$ and $\mathbb{T}_y = \{4, 10\}$, $k_x = k_2 = 2$, and f_x and f_y taking the average of their inputs. In this setting, the definition of $\mathbb{T}_x$ and $\mathbb{T}_y$ rules out time-acyclicity. The acyclicity assumption can only hold by effect-acyclicity, otherwise the variables cannot be validly represented in a causal DAG. Provided there is no causal cycle, the time order allows for any causal DAG between $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$. A causal DAG including the path $X_{\mathbb{T}_x} \rightsquigarrow Y_{\mathbb{T}_y}$ would require that $\mathbb{Y}_4$ does not cause $\mathbb{X}_6$, as in a non-adaptive medical treatment regime. Analogously, a causal DAG including the path $Y_{\mathbb{T}_y} \rightsquigarrow X_{\mathbb{T}_x}$ would require that $\mathbb{X}_0$ does not cause either of $\mathbb{Y}_4$ and $\mathbb{Y}_{10}$, and that $\mathbb{X}_6$ does not cause $\mathbb{Y}_{10}$.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

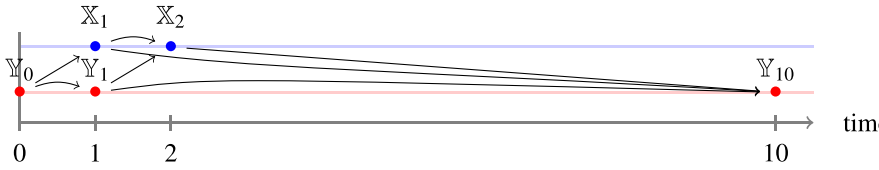


Figure 5. Atomic causal DAG.

Examples 1 to 3 also show why total effect-acyclicity is a strong assumption. In making a statement about all possible composite variables, total effect-acyclicity establishes acyclicity for all of the examples simultaneously. In settings with many interleaved time points this would likely be difficult, unless there is a fundamental asymmetry between the processes. Such an asymmetry exists, for example, between *solar flares* and *power outages*. The former can cause the latter, but never vice versa, making total effect-acyclicity plausible. However, this example also showcases that certainty about the absence of causal paths in one direction often comes at the cost of variables without physically well-defined interventions, in this case on the solar flares (for more cause–effect pairs, see Mooij et al. 2016). Finally, note that even given total effect-acyclicity, it is still necessary to specify the time of causal variables, since the edges still depend on the variables' temporal relationships.

4.1. Defining variables: Cycles, time scales, and unobserved confounding

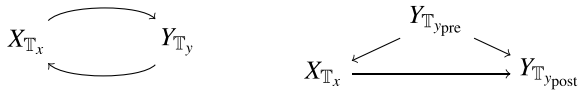
The issues of causal cycles, appropriate time scales, and unobserved confounding are of general interest to causal modelling, but they are particularly pertinent when the temporal nature of the variables is taken into account. Here, we illustrate how our formalism may be used to reason about them systematically.

Consider the interaction between the treatment *caffeine consumption* and the symptom *sleep* (cf. Stucky et al. 2024). Both variables are typically defined as aggregates over measurements at several time points, and hence cannot neatly be time-stamped with a single time point. Assume that, over a short period of time, sleep affects caffeine consumption and not vice versa, yet over a longer period, caffeine consumption also affects sleep. For illustration, let $(\mathbb{X}_t)_{t\in\mathscr{T}_x}$ be a stochastic process of sleep measurements (e.g. via brain activity) at time points $\mathscr{T}_x = \{1, 2\}$, and let $(\mathbb{Y}_t)_{t\in\mathscr{T}_y}$ be a stochastic process of caffeine measurements at time points $\mathscr{T}_y = \{0, 1, 10\}$. Let the atomic causal DAG be that shown in figure 5.

4.1.1. Unrolling causal cycles

It is often suggested that cyclic causal relationships are acyclic at a sufficiently fine-grained temporal resolution (e.g. Pearl 2009, sec. 1.4.1; Peters et al. 2017, sec. 2.3.3), meaning one could unroll cycles in time to obtain a DAG. Such an operation relies on a correspondence between variables and time points, which is what our formalism makes explicit. In the example, let for instance caffeine consumption be $X_{\mathbb{T}_x} := \mathbb{X}_2 - \mathbb{X}_1$, and sleep be $Y_{\mathbb{T}_y} := (\mathbb{Y}_0 + \mathbb{Y}_1 + \mathbb{Y}_{10})/3$. This gives rise to a cycle on the level of composite variables, as shown in figure 6(a). The causal cycle can be unrolled by splitting $Y_{\mathbb{T}_y}$ into the pre-treatment variable $Y_{\mathbb{T}_{y\text{pre}}} := (\mathbb{Y}_0 + \mathbb{Y}_1)/2$ and the post-treatment variable $\mathbb{Y}_{\mathbb{T}_{y\text{post}}} := \mathbb{Y}_{10}$. In combination with $X_{\mathbb{T}_x}$, these new

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press



(a) Cyclic causal graph. **(b)** Unrolled causal DAG.

Figure 6. Resolving a causal cycle by unrolling in time.

variables form the DAG in figure 6(b), which shows why the past of an outcome variable is a natural candidate for a confounder (cf. Imbens and Rubin 2015, sec. 1.8). Note that unrolling a causal cycle requires measurements at a higher temporal resolution and may change the interpretation of the model (cf. Hyttinen et al. 2012, sec. 1; Mooij et al. 2011, sec. 1). Composite variables in settings characterized by high-frequency interactions, such as stock prices or gene expressions, may therefore be difficult to unroll into meaningful DAGs.

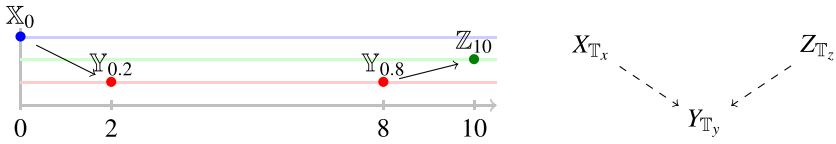
4.1.2. Time scales and unobserved confounding

Episodic settings such as those suggested by our medical examples provide a natural time scale, yet in general the time scale of causal models is not fixed and can affect their structure and properties (Weinberger 2020). For instance, in the example on caffeine consumption and sleep shown in figures 5 and 6, one could focus on a short time scale by reducing the model to $\mathbb{Y}_{tpre}$ and $X_{\mathbb{T}_x}$, or focus a long time scale by reducing the model to $X_{\mathbb{T}_x}$ and $\mathbb{Y}_{tpost}$. Though both of the resulting models are acyclic, the different time scales lead to opposite causal effects between the respective caffeine consumption and sleep variables. Moreover, in the long-term model there is confounding by $\mathbb{Y}_0$ and $\mathbb{Y}_1$, which highlights another important aspect of choosing an appropriate time scale. Quantities at time points that are not sampled may act as unobserved confounders, and the quantities modelled may exist even before the time horizon of the causal model, which greatly complicates justifying the absence of unobserved confounding. Though this has been noted in the context of time series (Thams et al. 2024), our temporal perspective on causal variables highlights that quantities at past time points may act as unobserved confounders in any causal setting.

4.2. Uncovering faithfulness violations

The foundational idea for discovering causal DAGs from data is to test for conditional independencies between variables to find a Bayesian network which may be interpreted as the causal DAG or equivalence class containing it (Spirtes et al. 2001). This approach relies upon the causal faithfulness assumption (see, e.g., Peters et al. 2017, definition 6.33). Using our formalism, we give an example to demonstrate that variables that refer to quantities at multiple time points may give rise to faithfulness violations that violate the time order. Faithfulness violations arising from variable groupings are described in (Parviainen and Kaski 2017, sec. 3) and further analyzed in (Wahl et al. 2024, secs. 5 and 6). This type of faithfulness violation cannot be easily dismissed by the arguments in Pearl (2009, sec. 2.4), Spirtes et al. (2001, sec. 3.5.2 and theorem 3.2), and Meek (1995b).

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press



(a) Atomic causal DAG. **(b)** Bayesian network of composite variables.

Figure 7. Faithfulness violations can lead to Bayesian network edges pointing backward in time.

Consider a variation of our earlier medical setting of treatment process $(\mathbb{X}_t)_{t\in\mathscr{T}_x}$ with $\mathscr{T}_x = \{0\}$, and symptom process $(\mathbb{Y}_t)_{t\in\mathscr{T}_y}$ with $\mathscr{T}_y = \{2, 8\}$. For this example, take the treatment to be aspirin and the symptom to be headache. Let further $(\mathbb{Z}_t)_{t\in\mathscr{T}_z}$ with $\mathscr{T}_z = \{10\}$ be a third process that measures absence from work. Let $X_{\mathbb{T}_x} := \mathbb{X}_0$, $Y_{\mathbb{T}_y} := (\mathbb{Y}_2 + \mathbb{Y}_8)/2$, and $Z_{\mathbb{T}_z} := \mathbb{Z}_{10}$. Assume further that aspirin at time 0 cures headache at time 2, and that headache at time 8 causes absence from work at time 10, as illustrated in the atomic causal DAG in figure 7(a). The temporal aggregation into composite variables leads to a pairwise causal dependence between $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$ and between $Y_{\mathbb{T}_y}$ and $Z_{\mathbb{T}_z}$, but not between $X_{\mathbb{T}_x}$ and $Z_{\mathbb{T}_z}$, meaning there is no transitivity. Applying the orientation rules in Meek (1995a) yields the Bayesian network shown in figure 7(b), which would be found using for example the PC algorithm for causal discovery (Spirtes et al. 2001, sec. 5.4.2). The edge $Z_{\mathbb{T}_z} \!-\!\!-\!\!\rightarrow Y_{\mathbb{T}_y}$ in the Bayesian network violates the time order.

This can be explained by a faithfulness violation. By definition A.2, the composite variables form the DAG $X_{\mathbb{T}_x} \rightarrow Y_{\mathbb{T}_y} \rightarrow Z_{\mathbb{T}_z}$. The variables $X_{\mathbb{T}_x}$ and $Z_{\mathbb{T}_z}$ are independent but not d-separated in that DAG. The example highlights that interpreting Bayesian network edges causally means reaching beyond their formal semantics, which requires careful justification (Dawid 2010, sec. 4.3).

5. Implications for statistical causality

The formalism presented in this paper aims to build upon the standard causal DAG framework. Nonetheless, our explicitly temporal composite variables mark a conceptual departure with far-reaching implications.

5.1. Causal discovery and causal representation learning

The field of causal discovery aims to infer causal graphs, usually DAGs or their equivalence classes, from data using statistical methods (see, e.g., Kitson et al. 2023). If one adopts our temporal perspective, this task suffers from an existential problem: Without prior knowledge about the causal order or the atomic DAG, how is one to know if an acyclic causal graph exists between the variables in the first place? For this reason, we consider the incorporation of time into causal DAGs a crucial step in identifying settings in which causal discovery can hope to succeed, and thus to address the lack of realism in causal discovery benchmarking (see, e.g., Gentzel et al. 2019; Reisach et al. 2021; Reisach et al. 2023; Brouillard et al. 2025). This implies a shift in perspective, away from discovering a causal order and toward treating it as necessary background knowledge (cf. Stern 2021; Bang and Didelez 2023). In the

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

context of learning causal representations and their corresponding DAGs from data (Schölkopf et al. 2021), our contribution highlights the importance of taking the time dimension of causal representations into account (cf. Weinberger 2020).

5.2. Causal effect estimation and causally robust prediction

A central benefit of causal DAGs is the do-calculus (Pearl 1995), which enables assessing identifiability and selecting suitable control variables for a causal effect of interest. However, the temporal relationship between variables may already provide useful information about suitable control variables, which explains why the potential outcomes framework enjoys great popularity for causal inference in the applied sciences (see, e.g., Angrist and Pischke 2009; Imbens and Rubin 2015), seemingly without a need for causal DAGs. Incorporating time into causal DAGs is likely to help outline the advantage they can offer more clearly. Taking time into account also provides a new perspective on the task of causally robust prediction (Bühlmann 2020), and the promise of causal DAGs for machine learning (Schölkopf 2022). In so far as variables correspond to single time points, causally robust prediction essentially amounts to predicting in the time order using quantities at the directly preceding causally relevant time points. However, many variables of interest may be composite variables with cyclic causal relationships, which could explain why the practical usefulness of causal machine learning remains to be demonstrated (Nastl and Hardt 2024).

5.3. The applicability of DAGs as causal models

Finally, and perhaps most importantly, taking into account the role of time means tackling the question of the applicability of causal DAGs to real-world data in earnest. Before the question *what is the right DAG?* comes the question *is there a DAG?* Treating causal variables as instantiations highlights the "statistical fantasy" (Spirtes and Glymour 1991, ch. 6) of assuming that all confounders have been observed, since this assumption may require accounting for past instantiations in an infinite regress. Moreover, the explicit role of time in our formalism clarifies that the validity of the acyclicity assumption is not self-evident. A justification via time-acyclicity requires an analysis of the measurement procedure, for example by using background knowledge or statistical inference to characterize the law governing measurement time points. A justification via effect-acyclicity requires an analysis of the causal relationships in the atomic DAG. Many variables of interest may not be amenable to causal analysis using DAGs, and what causal DAGs there are, are of a time-specific nature allowing for less sweeping conclusions. In this light, the crucial step in applying causal DAGs consists in defining meaningful variables that allow for their use, rather than treating a set of variables as given and reasoning about *the* DAG between them. Incorporating time into causal DAGs therefore offers the potential for a more pragmatic approach to causal inference in line with the practices of applied research.

6. Conclusion

Real-world quantities and their measurements exist and change in time. If causes are to precede effects, a realization of a causal model consists in a time series. It follows that causal relationships are relative with respect to the within-realization time

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

structure. For this reason, we argue that it is time (pun intended) to break with the tradition of nontemporal causal DAGs and incorporate time into the definition of variables in causal DAGs by default. This perspective reveals that causal DAGs do not relate generic concepts but temporal instantiations, which is essential for their interpretation and reconciles an important difference between causal analysis based on DAGs versus causal analysis based on potential outcomes. Variables that refer to quantities at multiple time points may give rise to causal cycles. In order to establish the validity of the acyclicity assumption, one must show that the variables refer to quantities at time points following an order that prevents causal cycles (time-acyclicity), or that there are no causal cycles if the time order allows them (effect-acyclicity). Incorporating time into causal DAGs is therefore essential for evaluating their applicability and interpretation as causal models. Our formal treatment of time allows for doing so in a principled fashion and opens many questions for future research. For example, it will be interesting to examine the change of causal effects with time distance, and to what extent some time points may be representative of others. Moreover, further research is required to analyze the implications for arguments regarding the prevalence of faithfulness violations, and to apply our notion of time to other classes of graphical causal models. Although our contribution shows the scope of causal DAGs to be more specific than it may otherwise appear, it also facilitates their discovery, clarifies their interpretation, and leaves them a less metaphysical and all the more practical tool for causal analysis.

Acknowledgments. We thank Niels Richard Hansen (University of Copenhagen) for helping to inspire this project, as well as Emilie Devijver (Université Grenoble Alpes), Pouya Babakhani (University of Bath), and especially Isabelle Drouet (Sorbonne Université) for helpful discussions and feedback throughout the stages of writing. AGR received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skodowska-Curie grant agreement No. 945332 and travel support under No. 951847, as well as mobility funding from Université Paris Cité via the 2024 BDMI program.

Funding Statement. None to declare.

Competing interests. None to declare.

References

Aalen, O. O., K. Røysland, J. M. Gran, R. Kouyos, and T. Lange. 2016. "Can We Believe DAGs? A Comment on the Relationship Between Causal DAGs and Mechanisms." *Statistical Methods in Medical Research* 25 (5):2294–314. https://doi.org/10.1177/0962280213520436.

Anand, Tara V., Adele H. Ribeiro, Jin Tian, and Elias Bareinboim. 2023. "Causal Effect Identification in Cluster DAGs. *Proceedings of the AAAI Conference on Artificial Intelligence* 37 (10):12172–79. https://doi.org/10.1609/aaai.v37i10.26435.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400829828.

Assaad, Charles. 2021. "Causal Discovery Between Time Series." PhD diss., Université Grenoble Alpes. https://hal.science/tel-03950986.

Bang, Christine W., and Vanessa Didelez. 2023. "Do We Become Wiser with Time? On Causal Equivalence with Tiered Background Knowledge." *Proceedings of Machine Learning Research* 216:119–29. https://proceedings.mlr.press/v216/bang23a.html.

Bongers, Stephan, Patrick Forré, Jonas Peters, and Joris M. Mooij. 2021. "Foundations of Structural Causal Models with Cycles and Latent Variables." *The Annals of Statistics* 49 (5):2885–915. https://doi.org/10.1214/21-AOS2064.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Brouillard, Philippe, Chandler Squires, Jonas Wahl, Konrad P. Kording, Karen Sachs, Alexandre Drouin, and Dhanya Sridhar. 2025. "The Landscape of Causal Discovery Data: Grounding Causal Discovery in Real-World Applications." *Preprint*. https://doi.org/10.48550/arXiv.2412.01953.

Bühlmann, Peter. 2020. "Invariance, Causality and Robustness." *Statistical Science* 35 (3):404–26. https://doi.org/10.1214/19-STS721.

Dawid, A. Philip. 2010. "Beware of the DAG! Causality: Abjectives and Assessment." *Proceedings of Machine Language Research* 6:59–86. https://proceedings.mlr.press/v6/dawid10a.html.

Dean, Thomas, and Keiji Kanazawa. 1989. "A Model for Reasoning About Persistence and Causation." *Computational Intelligence* 5 (2):142–50. https://doi.org/10.1111/j.1467-8640.1989.tb00324.x.

Gentzel, Amanda, Dan Garant, and David Jensen. 2019. "The Case for Evaluating Causal Models Using Interventional Measures and Empirical Data." In *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32. Available at: https://proceedings.neurips.cc/paper/2019/hash/a87c11b9100c608b7f8e98cfa316ff7b-Abstract.html.

Granger, C. W. J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (3):424–38. https://doi.org/10.2307/1912791.

Greenland, Sander. 2010. "Overthrowing the Tyranny of Null Hypotheses Hidden in Causal Diagrams." In *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, edited by Rina Dechter, Hector Geffner, and Joseph Y. Halpern, 365–82. College Publications.

Hansen, Niels, and Alexander Sokol. 2014. "Causal Interpretation of Stochastic Differential Equations." *Electronic Journal of Probability* 19:1–24. https://doi.org/10.1214/EJP.v19-2891.

Hernán, Miguel A., and James M. Robins. 2020. *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC.

Hill, Austin B. 1965. "The Environment and Disease: Association or Causation?" *Proceedings of the Royal Society of Medicine* 58 (5):295–300. https://doi.org/10.1177/003591576505800503.

Huang, Biewi, Kun Zhang, and Bernhard Schölkopf. 2015. "Identification of Time-Dependent Causal Model: A Gaussian Process Treatment." In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, edited by Qiang Yang and Michael Wooldridge, 3561. Palo Alto, CA: AAAI Press.

Hyttinen, Antti, Frederick Eberhardt, and Patrik O. Hoyer. 2012. "Learning Linear Cyclic Causal Models with Latent Variables." *Journal of Machine Learning Research* 13 (109):3387–439. https://jmlr.org/papers/v13/hyttinen12a.html.

Imbens, Guido W. and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139025751.

Kitson, Neville K., Anthony C. Constantinou, Zhigao Guo, Yang Liu, and Kiattikun Chobtham. 2023. "A Survey of Bayesian Network Structure Learning." *Artificial Intelligence Review* 56:8721–8814. https://doi.org/10.1007/s10462-022-10351-w.

Laplace, Pierre-Simon. 1814. *Essai philosophique sur les probabilités*. Paris: Courcier.

Lee, Hau-Chih, Matteo Danieletto, Riccardo Miotto, Sarah T. Cherng, and Joel T. Dudley. 2019. "Scaling Structural Learning with NO-BEARS to Infer Causal Transcriptome Networks." In *Biocomputing 2020*, edited by Russ B. Altman, A. Keith Dunker, Lawrence Hunter, Marylyn D. Ritchie, Tiffany Murray, and Teri E. Klein, 391–402. Singapore: World Scientific. https://doi.org/10.1142/9789811215636_0035.

Malinsky, Daniel, and Peter Spirtes. 2018. "Causal Structure Learning from Multivariate Time Series in Settings with Unmeasured Confounding." *Proceedings of Machine Language Research* 92:23–47. https://proceedings.mlr.press/v92/malinsky18a.html.

Meek, Christopher. 1995a. "Causal Inference and Causal Explanation with Background Knowledge." In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, edited by Philippe Besnard and Steve Hanks, 403–10. San Francisco, CA: Morgan Kaufmann Publishers. https://dl.acm.org/doi/10.5555/2074158.2074204.

Meek, Christopher. 1995b. "Strong Completeness and Faithfulness in Bayesian Networks." In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, edited by Philippe Besnard and Steve Hanks, 411–18. San Francisco, CA: Morgan Kaufmann Publishers. https://dl.acm.org/doi/10.5555/2074158.2074205.

Mooij, Joris M., Dominik Janzing, Tom Heskes, and Bernhard Schölkopf. 2011. "On Causal Discovery with Cyclic Additive Noise Models." In *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 24. Available at: https://proceedings.neurips.cc/paper/2011/hash/d61e4bbd6393c9111e6526ea173a7c8b-Abstract.html.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Mooij, Joris M., Jonas Peters, Dominik Janzing, Jakob Zscheischler, and Bernhard Schölkopf. 2016. "Distinguishing Cause from Effect Using Observational Data: Methods and Benchmarks." *Journal of Machine Learning Research* 17 (32):1–102. https://www.jmlr.org/papers/v17/14-518.html.

Murphy, Kevin P. 2002. "Dynamic Bayesian Networks: Representation, Inference and Learning." PhD diss., University of California, Berkeley. Available at: https://www.cs.ubc.ca/murphyk/Thesis/thesis.html.

Nastl, Vivian Y. and Moritz Hardt. 2024. "Do Causal Predictors Generalize Better to New Domains?" In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, edited by A. Globerson, L. Mackey, and D. Belgrave, 31202–315. Red Hook, NY: Curran Associates. https://dl.acm.org/doi/10.5555/3737916.3738899.

Parviainen, Pekka, and Samuel Kaski. 2017. "Learning Structures of Bayesian Networks for Variable Groups." *International Journal of Approximate Reasoning* 88:110–27. https://doi.org/10.1016/j.ijar.2017.05.006.

Pearl, Judea. 1995. "Causal Diagrams for Empirical Research." *Biometrika* 82 (4):669–88. https://doi.org/10.1093/biomet/82.4.669.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*, 1st ed. Cambridge: Cambridge University Press.

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803161.

Pearl, Judea, and Dana Mackenzie. 2018. *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books.

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. 2013. "Causal Inference on Time Series Using Restricted Structural Equation Models." In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, edited by C. J. C. Burges, L. Bottou, and M. Welling, 154–62. Red Hook, NY: Curran Associates. https://dl.acm.org/doi/10.5555/2999611.2999629.

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. 2017. *Elements of Causal Inference: Foundations and Learning Algorithms*. Cambridge, MA: MIT Press. Available at: https://library.oapen.org/handle/20.500.12657/26040.

Ramsey, Joseph, Madelyn Glymour, Ruben Sanchez-Romero, and Clark Glymour. 2017. "A Million Variables and More: The Fast Greedy Equivalence Search Algorithm for Learning High-Dimensional Graphical Causal Models, with an Application to Functional Magnetic Resonance Images." *International Journal of Data Science and Analytics* 3:121–29. https://doi.org/10.1007/s41060-016-0032-z.

Reisach, Alexander G., Christof Seiler, and Sebastian Weichwald. 2021. "Beware of the Simulated DAG! Causal Discovery Benchmarks May Be Easy to Game." In *Advances in Neural Information Processing Systems 34*, edited by M. Ranzato et al. Available at: https://proceedings.neurips.cc/paper/2021/hash/e987eff4a7c7b7e580d659feb6f60c1a-Abstract.html.

Reisach, Alexander G., Myriam Tami, Christof Seiler, Antoine Chambaz, and Sebastian Weichwald. 2023. "A Scale-Invariant Sorting Criterion to Find a Causal Order in Additive Noise Models." In *Advances in Neural Information Processing Systems 36*, edited by A. Oh et al. Available at: https://proceedings.neurips.cc/paper/2023/hash/027e86facfe7c1ea52ca1fca7bc1402b-Abstract.html.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5):688–701. https://doi.org/10.1037/h0037350.

Runge, Jakob. 2018. "Causal Network Reconstruction from Time Series: From Theoretical Assumptions to Practical Estimation." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28 (7):075310. https://doi.org/10.1063/1.5025050.

Schölkopf, Bernhard. 2022. "Causality for Machine Learning." In *Probabilistic and Causal Inference: The Works of Judea Pearl*, edited by Hector Geffner, Rina Dechter, and Joseph Y. Halpern, 765–804. New York: Association for Computing Machinery. https://doi.org/10.1145/3501714.3501755.

Schölkopf, Bernhard, Francesco Locatello, Stefan Bauer, Nan R. Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. 2021. "Toward Causal Representation Learning." *Proceedings of the IEEE* 109 (5):612–34. https://doi.org/10.1109/JPROC.2021.3058954.

Shoham, Yoav. 1990. "Nonmonotonic Reasoning and Causation." *Cognitive Science* 14 (2):213–52. https://doi.org/10.1207/s15516709cog1402_2.

Shpitser, Ilya, and Judea Pearl. 2008. "Complete Identification Methods for the Causal Hierarchy." *Journal of Machine Learning Research* 9:1941–79. https://jmlr.org/papers/v9/shpitser08a.html.

Spirtes, Peter, and Clark Glymour. 1991. "An Algorithm for Fast Recovery of Sparse Causal Graphs." *Social Science Computer Review.* 9 (1):62–72. https://doi.org/10.1177/089443939100900106.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Spirtes, Peter, Clark Glymour, and Richard Scheines. 1993. *Causation, Prediction, and Search*, 1st ed. (Lecture Notes in Statistics 81). New York: Springer. https://doi.org/10.1007/978-1-4612-2748-9.

Spirtes, Peter, Clark Glymour, and Richard Scheines. 2001. *Causation, Prediction, and Search*, 2nd ed. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/1754.001.0001.

Spirtes, Peter, Clark Glymour, Richard Scheines, Stuart Kauffman, Valerio Aimale, and Frank Wimberly. 2000. "Constructing Bayesian Network Models of Gene Expression Networks from Microarray Data." Available at: https://doi.org/10.1184/R1/6491291.

Spirtes, Peter, Clark Glymour, Richard Scheines, and Robert Tillman. 2010. "Automated Search for Causal Relations: Theory and Practice." In *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, edited by Rina Dechter, Hector Geffner, and Joseph Y. Halpern, 467–502. College Publications.

Splawa-Neyman, Jerzy. 1932. "On the Application of Probability Theory to Agricultural Experiments." *Statistical Science* 5 (4):465–80. https://doi.org/10.1214/ss/1177012031.

Spohn, Wolfgang. 2010. "The Structural Model and the Ranking Theoretic Approach to Causation: A Comparison." In *Heuristics, Probability and Causality: A Tribute to Judea Pearl*, edited by Rina Dechter, Hector Geffner, and Joseph Y. Halpern, 507–22. College Publications.

Spohn, Wolfgang. 2012. *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697502.001.0001.

Spohn, Wolfgang. 2019. "Reply to Jim Woodward's Comments on Wolfgang Spohn's Laws of Belief." *Philosophy of Science* 86 (4):773–84. https://doi.org/10.1086/704982.

Stern, Reuben. 2021. "Causal Concepts and Temporal Ordering." *Synthese* 198 (Suppl 27):6505–27. https://doi.org/10.1007/s11229-019-02235-4.

Stucky, Benjamin, Leonard Henckel, Marloes H. Maathuis, José Haba-Rubio, Pedro Marques-Vidal, Francesca Siclari, Raphaël Heinzer, and Hans-Peter Landolt. 2024. "Community-Based Causal Evidence that High Habitual Caffeine Consumption Alters Distinct Polysomnography-Derived Sleep Variables." *Journal of Psychopharmacology* 39 (12):1437–48. https://doi.org/10.1177/02698811251368364.

Suppes, Patrick. 1970. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland.

Thams, Nikolaj, Rikke Søndergaard, Sebastian Weichwald, and Jonas Peters. 2024. "Identifying Causal Effects Using Instrumental Time Series: Nuisance IV and Correcting for the Past." *Journal of Machine Learning Research* 25 (302):1–51. https://www.jmlr.org/papers/volume25/22-0262/22-0262.pdf.

Wahl, Jonas, Urmi Ninad, and Jakob Runge. 2024. "Foundations of Causal Discovery on Groups of Variables." *Journal of Causal Inference* 12 (1):20230041. https://doi.org/10.1515/jci-2023-0041.

Weinberger, Naftali. 2019. "Reintroducing Dynamics into Static Causal Models." In *Time and Causality Across the Sciences*, edited by Samantha Kleinberg, 187–211. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108592703.010.

Weinberger, Naftali. 2020. "Near-Decomposability and the Timescale Relativity of Causal Representations." *Philosophy of Science* 87 (5):841–56. https://doi.org/10.1086/710760.

Weinberger, Naftali. 2023. "Intervening and Letting Go: On the Adequacy of Equilibrium Causal Models." *Erkenntnis* 88 (6):2467–91. https://doi.org/10.1007/s10670-021-00463-0.

Williamson, Jon. 2004. *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198530794.001.0001.

Woodward, James. 2005. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press. https://doi.org/10.1093/0195155270.001.0001.

Wright, Sewall. 1934. "The Method of Path Coefficients." *The Annals of Mathematical Statistics* 5 (3): 161–215. https://doi.org/10.1214/aoms/1177732676.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Appendix
A. Formal definitions

In this section we provide additional details regarding our formalism and definitions introduced in Section 3. For convenience, we will use the following notation: for any $k \in \mathbb{N}^*$, for any set S of cardinality greater than or equal to k, $\mathbf{P}_k(S)$ denotes the collection of subsets of S which are of cardinality k. We denote the support of the distribution of $\mathbb{T}_x$ as $\mathbf{S}_x \subseteq \mathbf{P}_{k_x}(\mathscr{T}_x)$. We denote the support of the joint distribution of two sets of time points $\mathbb{T}_x$ and $\mathbb{T}_y$ with corresponding $\mathscr{T}_x$ and $\mathscr{T}_y$ as

$$\mathbf{S}_{xy} \subseteq \mathbf{S}_x \times \mathbf{S}_y \subseteq \mathbf{P}_{k_x}(\mathscr{T}_x) \times \mathbf{P}_{k_y}(\mathscr{T}_y). \tag{A1}$$

Furthermore, we denote the temporal precedence of $X_{\mathbb{T}_x}$ with respect to $Y_{\mathbb{T}_y}$ by the shorthand

$$\mathbb{T}_x \prec \mathbb{T}_y := \forall (s_x, s_y) \in \mathbf{S}_{xy}, \max(s_x) < \min(s_y). \tag{A2}$$

Definition A.1: Causes precede effects

We say that causes precede effects if and only if, in the atomic causal DAG between the temporal components of any pair of stochastic processes $(\mathbb{X}_t)_{t\in\mathscr{T}_x}$ and $(\mathbb{Y}_t)_{t\in\mathscr{T}_y}$,

$$\forall t_x \in \mathscr{T}_x, \forall t_y \in \mathscr{T}_y, (\mathbb{X}_{t_x} \rightsquigarrow \mathbb{Y}_{t_y} \Rightarrow t_x < t_y).$$

Definition A.2: Causation for composite variables

A variable $X_{\mathbb{T}_x}$ causes another causal variable $Y_{\mathbb{T}_y}$ if and only if there exists $(s_x, s_y) \in \mathbf{S}_{xy}$ with $t_x \in s_x$ and $t_y \in s_y$ such that $\mathbb{X}_{t_x}$ causes $\mathbb{Y}_{t_y}$ in the atomic DAG (we do not consider self-loops on the level of composite variables, e.g. if $\mathbb{X}_{t_1}$ causes $\mathbb{X}_{t_2}$ for some $t_1, t_2 \in s_x$ for $s_x \in \mathbf{S}_x$).

Definition A.3: Acyclicity

Acyclicity between two causal variables $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$ holds if and only if, in the atomic causal DAG,

$$\exists (s_x, s_y) \in \mathbf{S}_{xy}, \exists t_x \in s_x, \exists t_y \in s_y, (\mathbb{X}_{t_x} \rightsquigarrow \mathbb{Y}_{t_y}) \Rightarrow$$
$$\forall (s'_x, s'_y) \in \mathbf{S}_{xy}, \forall t'_x \in s'_x, \forall t'_y \in s'_y, \neg(\mathbb{Y}_{t'_y} \rightsquigarrow \mathbb{X}_{t'_x}).$$

We further call a graph $\mathscr{G}$ acyclic if acyclicity holds for all pairs of variables in $\mathscr{G}$.

Definition A.4: Time-acyclicity

Time-acyclicity between two causal variables $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$ holds if and only if

$$(\mathbb{T}_x \prec \mathbb{T}_y) \vee (\mathbb{T}_y \prec \mathbb{T}_x).$$

We further call a graph $\mathscr{G}$ time-acyclic if time-acyclicity holds for all pairs of variables in $\mathscr{G}$.

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press

Definition A.5: Effect-acyclicity

Effect-acyclicity holds between two causal variables $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$ if and only if there is no time-acyclicity by definition A.4, that is,

$$\neg((\mathbb{T}_x \prec \mathbb{T}_y) \vee (\mathbb{T}_y \prec \mathbb{T}_x)),$$

yet there is acyclicity by definition A.3. We further call a graph $\mathcal{G}$ effect-acyclic if effect-acyclicity holds for all pairs of variables in $\mathcal{G}$.

Definition A.6: Total effect-acyclicity

Total effect-acyclicity between two real-valued stochastic processes $(\mathbb{X}_t)_{t\in\mathcal{T}_x}$ and $(\mathbb{Y}_t)_{t\in\mathcal{T}_y}$ holds if and only if, in the atomic causal DAG,

$$\exists(t_x, t_y) \in \mathcal{T}_x \times \mathcal{T}_y, (\mathbb{X}_{t_x} \rightsquigarrow \mathbb{Y}_{t_y}) \Rightarrow \forall(t'_x, t'_y) \in \mathcal{T}_x \times \mathcal{T}_y, \neg(\mathbb{Y}_{t'_y} \rightsquigarrow \mathbb{X}_{t'_x}).$$

In this case, we say that total effect-acyclicity holds for any $X_{\mathbb{T}_x}$ and $Y_{\mathbb{T}_y}$ defined with respect to $(\mathbb{X}_t)_{t\in\mathcal{T}_x}$ and $(\mathbb{Y}_t)_{t\in\mathcal{T}_y}$, respectively. We further call a graph $\mathcal{G}$ totally effect-acyclic if effect-acyclicity holds for all pairs of variables in $\mathcal{G}$.

Cite this article: Reisach, Alexander Gilbert, Alberto Suárez, Sebastian Weichwald, and Antoine Chambaz. 2026. "The Case For Time in Causal DAGs." *Philosophy of Science*. https://doi.org/10.1017/psa.2026.10216

https://doi.org/10.1017/psa.2026.10216 Published online by Cambridge University Press