Forthcoming; manuscript No.
(will be inserted by the editor)

On the Ramifications of Theory Choice in Causal Assessment

Indicators of Causation and Their Conceptual Relationships

Roland Poellinger

Received: date / Accepted: date

Abstract Causal reasoning can be seen as the backbone of science in aiming at the attainment of stable, transferable, and replicable knowledge. In the empirical sciences, gathering data helps the researcher in forming an opinion about the investigated causal hypothesis. What makes collected data evidence for or against such a hypothesis is its interpretation in light of the indicative traces the hypothetical causal association entails. In this paper, I will first collect the pivotal causal indicators. I will then revisit the literature on causation and put on different hats in reassessing how theory choice influences the relationships amongst this set of indicators. I will finally argue that, for the precise formulations proposed, utilizing all four indicators (i) presents a resolution to seemingly conflicting perspectives in the causal literature, and (ii) paves the way towards an enriched picture of scientific instruction, prediction, and explanation across the conceptual fences of different causal theories.

Keywords Epistemology and metaphysics of causation, theories of causation, causal inference

Contents

This work is supported by the European Research Council (grant 639276) and the Munich Center for Mathematical Philosophy (MCMP).

Roland Poellinger
Munich Center for Mathematical Philosophy, LMU Munich
E-mail: r.poellinger@lmu.de

1 On the track of a causal association

Causal reasoning can be seen as the backbone of science in aiming at the attainment of stable, transferable, and replicable knowledge. In the empirical sciences, gathering data helps the researcher in forming an opinion about the investigated causal hypothesis which is rarely tested directly. What makes collected data evidence for or against the hypothesis is their interpretation in light of the indicative traces the hypothesized causal association entails: The more such empirical data supports the theoretical consequences (i.e., testable semantical aspects) of a causal hypothesis, the higher our confidence in the existence of the causal association, precisely because these theoretical consequences speak to the truth of the hypothesis as mediating *indicators*. In this paper, I will first collect the pivotal causal indicators and then discuss their interdependencies through the lenses of the most prominent theories of causation. I will then advance the view that making these indicators precise and keeping them distinct facilitates a unifying view on seemingly contradictory theoretical stances and thus paves the way for a novel, cross-theoretical concept of causal assessment. But let me start by carving out the pivotal indicators of a causal association.

The philosophical literature on the epistemology and ontology of causality distinguishes three tightly related main goals of the hunt for causal knowledge: *control*, *prediction*, and *explanation*. The idea that causal knowledge bestows control on us is rooted in the intuition that causes are difference-makers and can thus be leveraged to bring about desired effects (or influence expected effect events toward a desired direction, respectively). Possessing knowledge of such control factors allows us to instruct others in their pursuit of specific outcomes. Control can be understood as a special instance of prediction: Predicting the effects of my own actions precisely enables me to choose the action suited to bring about the predicted effect. When it comes to predicting outcomes whose causal factors are not (or not fully) under our control though, we rather couch our statements about suspected effects in a language that conveys our uncertainty about possibly complex scenarios. We are prepared to encounter exceptions in the chain of causal links and oftentimes resort to speaking of tendency and likeliness: a statement such as "smoking causes lung cancer" is to be understood as a comparative claim about risk, i.e., ultimately, a claim about probabilities (and how these change). Knowledge about difference-making and information about tendencies may both figure in the third main aim of causal reasoning, explanations. Nevertheless, if a causal explanation is given in an attempt to satisfy a question about how exactly it was that a specific effect occurred, the explanation will include semantically rich details about the workings of the system under consideration.

Instructions, predictions, and explanations are fundamentally grounded in time – they will only be fruitful and successful in a given context if they do not run counter to how the causal and the temporal structure of Nature are related. If causes precede their effects, then they will make a difference to future events, they will shape temporally posterior probabilities, and they will figure as genuinely initial conditions for fine-grained descriptions of a system's interacting components in informative answers to

Theory	D-M	PROB	MECH	TEMP
Hume I		✓		✓
Hume II	✓			(✓)
Lewis	✓			(✓)
Mackie	✓			(✓)
Reichenbach		✓		✓
Suppes		✓		✓
Good		✓		✓
Spohn		✓		✓
Pearl	✓	✓		
SGS	✓	✓		
Woodward	✓	✓		
Salmon	✓		✓	
Dowe	✓		✓	
Popper		✓	✓	
RWT		✓	✓	

Table 1: Systematic overview over how various prominent theories of causation build on different indicators of a causal relation in their explication of cause and effect.

how-questions. Virtually no theory of causation stays silent on how its concepts relate to time: If time does not explicitly figure in the definiens of a causal relation, the direction of causation is stipulated as along the arrow of time, because temporal order plays a crucial role for the identifiability of cause and effect – as deeply entrenched in our intuitions, as established by successful heuristics, or as encoded in physic's differential equations (possible deviations from such alignment have been philosophically motivated and defended; see, e.g., [9] and [32]).

When empirical data informs us about the difference-making or probability-raising relation between cause and effect, or about a system's mechanistic and temporal structure, these semantical aspects point to the truth (or falsity) of a hypothesized causal association as testable *causal indicators*. Fig. 1 provides a systematic overview over how different conceptual approaches towards causation build on these four pivotal indicators of a causal association in their explication of cause and effect, with *D-M* as label for the difference-making aspect of a cause, *PROB* referring to both characteristics of probabilistic covariation, dependency and uncertainty, *MECH* as label for knowledge about the connecting mechanism supporting the causal association, and *TEMP* as label for information about the temporal order of cause and effect:

1. *Hume I* refers to the first half of Hume's famous dictum: "We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second." ([14, Section VII]). Building on constant conjunction, this analytic departure from a purely metaphysical concept

of causation can be seen as an important starting point for regularity theories of causation.

2. *Hume II* refers to the Hume's *other words*, the second half of his famous definition: "[I]f the first object had not been, the second never had existed." ([14, Section VII]; it might be argued about whether *first* and *second* have a temporal meaning here). This can both be understood as the starting point for counterfactual as well as for interventionist theories of causation, as listed in the following.
3. *Lewis* actually picked up on Hume's second half of the quote and framed his concept of causation in terms of counterfactuals (to be evaluated in possible worlds). In his first formulation ([19]), Lewis does without time and only builds on the asymmetry of difference-making ('small miracles' disturbing the normal course of events). In further elaborations ([20]), Lewis introduces time in an attempt to avoid overdetermination via fine-graining.
4. *Mackie's* famous INUS condition goes under the label 'logical reconstruction' – in this explication, a cause is defined as a particular element in a logical schema by which the relevance of a cause to its effect (D-M) is captured in terms of necessity and sufficiency. It shall be mentioned that *Streven's* kairetic account is formulated in the same vein (see [40]). Although Mackie does not include time in his definition, he discusses the spatio-temporal distance and order of the INUS relata explicitly (see, e.g., *Causes and Conditions*, [22, p. 258]).
5. The names of *Reichenbach* ([33]), *Suppes* ([41]), *Good* ([12]), and *Spohn* ([38]) are all connected with theories integrating probabilities and time in a Humean effort to learn about the causal structure of the world only from observational data.
6. *Pearl* (together with Verma and Halpern, see [27], [28], [42], [13], [13]), *Spirtes, Glymour, and Scheines (SGS)* ([37]), as well as *Woodward* ([44]) explicate causality in terms of probabilistic concepts, and build on the Bayes net framework to accommodate the idea of asymmetrical difference-making in the sense of system-external interventions (distinguished from personal agency).
7. *Salmon* ([35]) and *Dowe* ([8]) are best known for emphasizing the push-and-pull mechanism (MECH) between cause and effect, in terms of traceable physical processes or exchange of conserved quantities. Although the proponents of these ideas aimed at freeing the concept of causal relations from non-physical knowledge, implicitly, this move cannot do without counterfactual information (D-M) about how things would have been had two processes not interacted, or had energy, momentum, etc. not been transferred.[1]
8. *Popper* argues for an objective interpretation of conditional probability capturing causally productive disposition as "a property of the generating circumstances" ([31, p. 34]). In Popper's view, probabilities "must be physical propensities, abstract relational properties of the physical situation, like Newtonian forces, and 'real', not only in the sense that they could influence the experimental results, but also in the sense that they could, under certain circumstances (coherence), interfere, i.e. interact, with one another" ([31, p. 28]). Causal theories in that spirit,

[1] [8] notes that "Salmon was uncomfortable with the use of counterfactuals, fearing that their context dependence was problematic for an 'ontic' account.", and examines Salmon's stance ([35]) in light of Kitcher's ([17]) critical comments.

in one way or another, combine probabilistic knowledge about causal tendencies (PROB) and a process or systems view (MECH) to explicate productive capacity or causal power (see also, e.g., [23] and [24]).

9. I want to consider one more position in this list: *RWT*, the *Russo–Williamson Thesis*, is highly debated in the literature on causation in medicine and can be understood as in opposition to causal inference based on merely probabilistic data (as advocated in the evidence-based medicine paradigm, EBM). RWT suggests that only both an appropriate correlation and mechanistic knowledge together genuinely establish a causal claim (see [34]).

This list is certainly not exhaustive, but it illustrates how the theoretical differences between prominent analytic theories of causation can be broken down in terms of the four indicators introduced above. Notably, all of the theories discussed (explicitly or implicitly) rely on a combination of two aspects.

My plan for this paper is to first make the content of the four causal indicators introduced above explicit. After doing this I will revisit the literature on causation and put on different hats in reassessing how theory choice influences the relationships among this set of indicators, and how each indicator relates to the hypothesized causal association. I will then argue that, for the precise formulations proposed, utilizing all four indicators (i) presents a resolution to seemingly conflicting perspectives in the causal literature, and (ii) paves the way towards an enriched picture of scientific instruction, prediction, and explanation across the conceptual fences of different causal theories.

2 Evaluating causal indicators with different causal theories

Before I want to do a comparative survey of how the conceptual relations between our four indicators change with respect to one's choice of causal theory, I will explicate the content of our indicators on a level of abstraction suitable for the present purpose. In doing so, I am not aiming at establishing a logical system as a frame for a rigid proof. My goal is to delineate the indicators in such a way that I can juxtapose them on very much the same analytic level with different theoretical explications of cause and effect.

The four causal indicators introduced above encode facts about the investigated causal association. Let me list the four indicators again together with their shorthand slogans and more elaborate explications in which I am relating intrinsic connotations and sub-properties of their respective relational characteristics:

D-M: Causes make a difference to their effects; or more elaborate: A cause event makes a difference (influence) or can be used to make a difference (control) to the normal course of events (as counterfactual deviant) in bringing about its effects.
PROB: Causes and effects covary; or more elaborate: A cause event will change the probability of its cause (dependency) within a particular expectancy (predictive uncertainty).

MECH: Causes and effects are connected by a mechanism; or more elaborate: There exists a mechanism by and through whose interconnected and interacting components the cause influences its effect.
TEMP: Boldly put, causes precede their effects; or a bit more elaborate and a bit more cautious at the same time: cause and effect are in a temporal relation (w.r.t. order, duration, and distance) that is compatible with one's understanding of how time and causation hang together.

It is important to remark here that these indicators are of different strength in causal assessment: TEMP might only be less informative yet a necessary requirement, while D-M might be understood as a perfect indicator (i.e., a sufficient condition) *entailing* the existence of a causal association.

The following four sections will look in more detail at how our four indicators are spelled out in different causal theories, and how their conceptual relations vary for these different causal theories. To this end, I will revisit D-M (Subsec. 2.1), PROB (Subsec. 2.2), MECH (Subsec. 2.3), and TEMP (Subsec. 2.4) by putting on different 'theoretical hats' (not necessarily all hats in the list above for all points of comparison, but in each case at least one pair of hats that provoke contrasting viewpoints). To structure the text, I will highlight the discussion of such different hats by marking the theoretical approaches listed in Tab. 1 in bold face. For illustrative purposes I will use some of the standard examples found in the literature together with examples taken from medicine, where the analysis of cause and effect comes with notorious difficulties and the theoretical/conceptual frame for causal assessment is far from settled.

2.1 Causes make a difference to their effects (D-M)

*D-M and PROB.*Keeping the causal indicators D-M and PROB separate reflects the conceptual/methodological dividing line between observational/static and interventional/dynamic support for the causal hypothesis, i.e., the opposition between inference from observation alone vs. inference from data collected in interaction with the investigated system or population. As mentioned above, this principled distinction is already laid out in Hume's famous twofold definition of causation which can be seen as a point of reference both for **regularity/supervenience theories** as well as for **counterfactual/manipulationist theories** of causation (cf. [14, Sec. VII]). Proponents of the latter camp emphasize in their theories the asymmetrically directed, difference-making relationship between cause and effect (in contrast to symmetric regularities; see also [18] for a discussion of difference-making characteristics). For example, in Pearl's **interventionist account of causation**, D-M and PROB surface in the mathematical explication of cause and effect as two distinct theoretical components: the difference-making power of a cause is mirrored in changed outcome values upon symmetry-breaking transformation of the investigated system, whereas uncertainty is expressed as disturbance terms inside the structural equations inducing the topology of the causal network. Causal mechanistic functions of the form $f_i(\mathbf{pa}_i, u_i)$

(determining *in the observational case* the outcome of effect variable V_i w.r.t. its direct causes – i.e., parents in the network – $\mathbf{pa}_i$ and the sum of potential random disturbances u_i) are replaced by constant values to express the forced value of V_i upon intervention on the system. This symmetry-breaking intervention is mirrored in the graph as removal of edges from V_i's parents into V_i. Thus, what difference-making via intervention means is really defined over a *pair* of Bayesian networks (for a formal statement of this contrast, see especially Pearl's definition of a probabilistic causal model, Def. 7.1.6 in [28, p. 205], and his explication of causal claims in terms of structure-based counterfactuals, [28, pp. 222 f.]).

From the perspective of the empirical sciences, the dividing line between D-M and PROB becomes blurred: When the **Russo-Williamson Thesis (RWT)** ([34]) calls for the combination of *difference-making* plus *mechanistic knowledge*, what is really meant is *appropriate correlations* plus *biologically plausible explicatory system descriptions* (in order to improve upon the merely data-driven orientation the evidence-based medicine paradigm advocates). RCTs, for example, have been designed to yield robust difference-making information about an investigated system. Yet, even when diligently devised, it is simply never possible to consider (let alone, measure) all possible confounding factors potentially influencing the investigated causal association. The "ideal RCT", in which the investigated causal association is genuinely isolated from other influences, must be understood as a guide in the scientific endeavor. Consequently, if one skeptically claims that only the unattainable, ideal RCT yields genuine difference-making information (along the lines of [45]), then one has to accept that any real-life experiment only returns correlations (dependency and strength) – some more, some less appropriate.

*D-M and MECH.*I will treat the relationship between the causal indicators of difference-making and mechanistic connection below after subjecting MECH to close scrutiny in Sec. 2.3.

*D-M and TEMP.*The counterfactual-based cases of Hume II, Lewis, and Mackie suggest that **counterfactual intuitions** (as one way of spelling out difference-making) better be supplemented with temporal information to avoid certain causal pitfalls: E.g., so-called *backtracking conditionals* must be ruled out by additional knowledge if one wants to maintain the intuition that causes precede their effects. The cluster of **interventionist and process accounts** in the lower half of Tab. 1, however, shows that the temporality aspect of a hypothesized causal relation seems to be secondary if mechanistic details are available or the control aspect of D-M is emphasized – quite naturally so: (i) Mechanistic information is understood to provide insights into a system's development *over time*, and (ii) knowledge about how to control a system can be used in *predicting* the outcome of one's own actions (or, similarly, in instructing someone else). Pearl ([28]), SGS ([37]), and Woodward ([44]) augment their probabilistic frameworks in a way such that causally interpreted conditional probabilities only allow for causal flow in one direction.

*D-M and the causal claim.*As mentioned above, D-M can be understood as a *perfect indicator* of a causal relation. But, to qualify this statement: Although in the

empirically-informed literature on causal inference *Randomized Controlled Trials* (RCTs) are advocated as the means of choice for detecting difference-makers, any RCT is necessarily bound to a set of variables deemed relevant and sufficient for the task. The output of an RCT will thus only yield partial information about D-M in the idealized absolute sense. Consequently, even though D-M might be a perfect indicator of a causal relation, there is simply no perfect data to match this indicator. We have to keep this in mind going forward (see, e.g., [45] and the paragraph on *D-M and PROB* above.).

Moreover, even if one understands D-M as a perfect indicator of a causal relation, it is not a conceptual necessity to infer from the truth of a causal claim the possibility of control via difference-making. Another look into the discussion around the efficacy and validity of RCTs reveals why. In their discussion of causal inference from RCTs, Landes, Osimani, and Poellinger ([18]) state the following:[2]

> Methodological pluralists such as Cartwright [7,5], and Stegenga [39], among others, express concerns against the privileged role of RCTs also on grounds that classical 'linear' approaches to causal inference cannot do justice to the complexity of causal phenomena in the biological and social sciences, characterized by nonlinear causation and causal interactions. In the same line, also modular conceptualization of causes such as the ones implied in the causal graph methodology developed by Pearl [28] and Glymour [37] and colleagues (see also [44]), are under attack for failing to recognize that causes may be holistic and therefore may be not adequately captured by a difference making account.

Keeping a causal relation and its indicator D-M separate leaves room for **holistic causation** in systems that cannot be investigated with the modularity-based interventionist toolbox. Pearl as prominent proponent of the **interventionist account of causation** rejects the idea of holistic causation as a contradiction in terms. Not only does he define what it means to be a cause based on atomic, surgical interventions, he also emphasizes repeatedly that any practical task that can be described in causal vocabulary can also be cast in a causal model and consequently precisely be specified and analyzed in terms of hypothetical interventions (see in particular Pearl's reply to Cartwright's critique in Appendix 11.4.6 of [29, pp. 362 ff.]). From this perspective (and contrary to the holist's position), knowing that the causal association holds is equivalent to knowing that this relationship is characterized by the difference-making property.

Now, one could either say that Pearl's account simply rules out the idea of holistic causation precisely in tying causation to the possibility of interventions, or that his account only covers a subset of causal scenarios – namely those that can be analyzed via intervention. Yet, even for the seemingly irreconcilable clash between Pearl and Cartwright on this matter, a conciliatory middle-ground can be found in Woodward's relativization of causal efficacy to the question whether a suitable intervention

[2] See also Osimani's paper on the asymmetry of evidence for risk [25] as well as the concerns expressed in [6], [24], [1], and [16]

variable can *practically* be defined. In this move, Woodward ([44]) separates causal structure (in the same Bayes-net spirit as Pearl) from interventions which might not be straightforwardly guaranteed (in contrast with the abstractly defined, purely structural transformations in Pearl's proposal). In this sense, Woodward combines Cartwright's and Pearl's analysis in his fine-grained sub-case definition of causal effect. His account can be understood as one further, illuminating example of how collecting different understandings of/stances towards difference-making, encoded in different causal theories can be used in strengthening prediction, explanation, and instruction (control) – even if (or: even more so if) the respective theories are seemingly in conflict. Key to this enriched concept of cross-theoretical causal assessment is the distinction between the hypothesis and its indicators – and the distinction between different indicators of causation. I will now turn to the causal indicator PROB and investigate its role in contributing to enriched, cross-theoretical causal assessment.

2.2 Causes and effects covary (PROB)

We already discussed the relationship between PROB and D-M above (Sec. 2.1); how PROB and MECH are related shall be looked at in detail in the next section (Sec. 2.3) – so we are left with the question how PROB and TEMP are related, and if their relationship varies for different 'theoretical hats'.

*PROB and TEMP.*The relationship between probabilistic and temporal information is really at the heart of much of the debate on both the ontology and the learnability of causality since Hume (*Hume I* in the table above). Reichenbach's common cause principle ([33]) highlights the necessity of temporal information in learning about the direction of a causal association: Correlation only tells us *that* some causal structure supports the dependency, but not *which*. In **probabilistic theories of causation**, information about dependency and/or tendency must thus be supplemented with temporal information to discern sound models from misrepresentative ones. Notably, Popper deviates from this requirement in his formulation of the propensity account – he invokes a capacity understanding of a causal system in order to treat quantum experiments in which precisely the idea of a cause's temporal precedence seems to be violated. The proponents of the **interventionist account of causation** do not include time in their explications, but – as mentioned above (in the paragraph on D-M and TEMP) – temporal information becomes secondary once knowledge about control (difference-making via intervention) is available. One important thing shall be mentioned here, nevertheless: When the interventionist account of causation is implemented to infer causal relations from static, non-interventional data with causal learning algorithms (e.g., with Pearl's algorithm for *Inferred Causation*, *IC**), time comes back into the picture explicitly as one way of narrowing down the class of *observationally equivalent* (*Markov-equivalent*) causal structures. Acquiring information about the temporal structure of a causal association is not always straightforward nor trivial, since temporal order, distance, and duration may interact such that variable choice (or definition) becomes an arduous task. For example, in epidemiology, where the onset of a certain disease might not have been detected or well documented, the

underspecified duration of the disease might make it difficult to avert confounding by indication.

*PROB and the causal claim.*I want to conclude this subsection with a some brief remarks on the relationship between the indicator PROB and the hypothesized causal association: According to Reichenbach's principle, there is no correlation without causation; i.e., when PROB points to probabilistic dependency between cause and effect candidate, this precisely means that they stand in *some* causal relation (may it be a common cause structure). Reichenbach's principle together with the possibility of capturing uncertainty with random variables has been exploited by probabilistic approaches in order to deliver better *predictions* by simulating exceptional influences and unmodeled disturbances. Yet, problems emerge when it comes to extracting causal associations from our surroundings: Firstly, concrete real-life data will in all cases show dependencies and tendencies between any two variables. The negative signal 'independency/no tendency' must obviously be relativized to some threshold, i.e., in statistical terms, some arbitrarily or habitually fixed significance level (in other words: algorithms for the detection of independencies are only as good as the power of the statistical test they encode). And secondly, causation does not entail correlation, in reverse (see especially Cartwright's critical comment 'What is Wrong with Bayes Nets?', [3]):

(i) Causes do not necessarily raise the probability (tendency) of their effects (e.g., a badly hit golf ball might arrive at the hole only because some tree's branch, usually an obstacle, diverted the ball in the right direction),
(ii) in Simpson's paradox it depends on the choice of population whether some variable raises or lowers the probability of a second dependent one,
(iii) causal dependency might be hidden by statistical independence through precise cancellation (as, e.g., in Hesslow's famous pill example),
(iv) same time trends may be statistically dependent, but not causally linked (as, e.g., when the bread price in London and the sea level in Venice rise alike), and
(v) correlation (or covariation, respectively) might simply be due to other reasons like logical, mathematical, semantical dependencies (see also [43]).

Obviously, probabilistic causal analysis is well-advised to include evidence for a second causal indicator in order to avoid such pitfalls. Table 1 shows how different theories emphasize additional concepts – with different advantages and different disadvantages (for the combination of PROB + DM see Sec. 2.1, for the combination of PROB + MECH see Sec. 2.3, and for the combination of PROB + TEMP see Sec. 2.4 below).

With current achievements in statistical data analysis and machine learning techniques, causal inference from even large pools of data has become attainable. This opens up the possibility of synthesizing big and unstructured, heterogeneous and also historical data. Reichenbach's principle tells us that PROB alone will not be a good idea for causal analysis, and causal learning algorithms must be supplemented with further information to narrow down the output class. Viewing causal assessment from

the here proposed higher perspective in distinguishing the indicators of a causal hypothesis makes transparent how automated causal inference techniques can benefit from cross-theoretical causal assessment – in specifying heuristics, in encoding expert knowledge, and in defining points of input for supervised learning strategies.

2.3 Causes and effects are connected by a mechanism (MECH)

The concept of mechanism is highly debated in the philosophy of science, with different characterizations emphasizing different qualitative aspects of what a mechanism could be and what the concept might be useful for.[3] However, a common denominator of this enterprise seems to be that mechanistic reasoning is invoked in contrast to black-box inference, precisely when questions about the innards of the black box arise. Therefore, one idea common to different conceptualizations is modularity: A mechanism is somehow composed of different parts that together fulfill a certain function, may it be to propagate something, to sustain itself or a larger system, or also to act as a switch for other connecting mechanisms in relaying inputs to different outputs. While the causal indicator MECH, as introduced above, refers to the complete mechanism, in biology or pharmacology for example it is usually the case that evidential support for MECH comes in small pieces (about sub-mechanisms). How such piecemeal evidence is glued together in support of MECH and what it means for MECH to be partially "confirmed" are difficult questions in themselves and go beyond the scope of this paper. As for D-M and PROB, I want to investigate MECH's relation to the other indicators from different theoretical perspectives in the following – yet, what could be our handle on this elusive creature?

MECH was introduced above as encoding the existence of a mechanism *by and through whose interconnected and interacting components the cause influences its effect*. This existential claim is easy prey to charges of vacuousness, if one understands Nature as the all-embracing mechanism whose interconnected and interacting components support causation around and within us. To avoid this conceptual erosion I want to fully embrace the fact, that the indicator MECH is *perspectival* in character in the following sense: (i) Mechanistic reasoning is invoked in going beyond black-box inference, when how-questions arise and the need for a stable qualitative description of a system's inner workings emerges – such contexts establish *perspectivally* if and when an answer is deemed satisfying, and when a description is good enough to support explanation. (ii) Depending on the investigation's level of abstraction, the concept of mechanism will possess a certain degree of definedness – from strictly defined technical term to loosely defined but in a given context well-understood pragmatic label. And finally, (iii) if the existence of a mechanism is to be checked recursively by going through its constituents, it depends on the level of (descriptive) granularity how the existential claim can be supported or refuted.

[3] See, e.g., [21] as a point of reference for much of the current debate, and [10] as well as [11] for a more recent overview.

Embracing this perspectival character of the mechanism concept, I will couch the discussion in this section in terms of knowledge about the mechanism, i.e., I will understand a mechanism's existence in terms of the availability of a qualitatively rich, fine-grained, explicatory description of a mechanism's components, their connections, and their interactions.

MECH and D-M. **Causal process theories** offer the most tangible interpretation of MECH in terms of fundamental connections: Salmon, for example, famously defines a causal process based on the concept of mark transmission (see, e.g., [8]). A billiard ball uniformly moving across a billiard table can be understood as a causal process, because the ball itself has the capacity of transmitting a physical mark applied to it, e.g., a stroke of chalk. In contrast, a shadow moving across the same billiard table has no causal powers in that sense, because it cannot be marked physically in the same way. This idea is also very close to the medical practice of using chemical tracers to map biological pathways. But in itself such practice only hints at the locus of *possible* interactions and of *possible* ways of difference-making, and it does not reveal the inner workings of how effects-as-differences are actually brought about. Much criticism has been directed towards process theories in this vein, precisely because Salmon's definition of causal interaction relies on the counterfactual aspect of D-M, i.e., knowledge about how things would have been – exactly the type of knowledge, the idea of physical mark transmission sought to circumvent.

In a refined theory of causal processes, Dowe ([8]) draws on the concept of conserved quantities (like charge, momentum, etc.) in his definition of a causal interaction: causal processes (possessing a conserved quantity) interact causally by exchanging conserved quantities (at the crossing of their trajectories). These attempts at describing what it means to be a causally efficacious event were really meant to capture a physics-based observable/measurable and replicable comprehension of a causal connection. When theoretically scrutinized, though, these attempts might yield too many causes (e.g., causes at different levels), and require knowledge about the unperturbed course of events – so to speak, a *counterfactual*, thus per definitionem not observable, baseline.

Furthermore, although causal processes and exchange of conserved quantities seems so intuitive, solid theories for complex systems are lacking. Nevertheless, describing causal processes as difference-makers is possible, if the description of the mechanistic relationships includes the causal process, the level at which it operates, the locus of the interaction, and the quantity to be considered together with the description of the unperturbed system. Yet, as soon as all these ingredients are available, we are essentially back to the interventionist concept of causation (which the causal process theorist tried to avoid in the first place), and MECH is demoted to a conceptual part of D-M (e.g., as *difference-making at a certain biochemical level of description*).

Interventionist causal theories provide an explication of the mechanism concept by building on the control aspect of D-M: In contrast with the rather family-like, fuzzy, pragmatics-infused concept of mechanism above, Pearl ([28]), for example, rigorously *defines* a causal mechanism to be a function f_X, assigning to a random

variable X (representing an event) its value x by computing only its direct causes, i.e., X's parent variables (w.r.t. a given ordering in a faithful Bayes net causal model). Pearl goes on to define C to be a cause of E iff C and E are probabilistically dependent upon intervening on the causal structure: "Manipulation subjugates the putative causal event to the sole influence of a known mechanism, thus overruling the influence of uncontrolled factors that might also produce the putative effect" ([28, p. 253]). While C is lifted from the influence of its parents, the only way for C to influence E is along a web of causal mechanisms supporting the path from C to E. For Pearl, it does not make sense to speak of a difference-making relationship without a mechanism underpinning it, and it does not make sense to define or postulate a mechanism that cannot be intervened on. In this sense, the causal indicators D-M and MECH mutually depend on each other and cannot be separated. One important consequence must be mentioned here: If one completes the reductive move and postulates D-M as exhaustive characterization of causation (see the above discussion about the relation between D-M and the causal claim), then MECH and D-M really disappear as indicators but much rather present modes of piecemeal causal learning.

This move comes with a caveat, though: Collapsing D-M, MECH, and the causal association into one conceptual building block might be legitimate from an interventionist's perspective, if the causal hypothesis speaks about the same domain (population, scenario, etc.) as the evidence that supports the indicators D-M and MECH. If, in contrast, D-M and MECH are supported by evidence from a domain (population, scenario, etc.) that differs from the intended target the causal hypothesis is meant to apply to, then D-M plus MECH and the investigated causal association should be kept apart. Bareinboim and Pearl (2012) discuss criteria for the transportability of difference-making knowledge from experimental results to distinct targets for fixed structures. But things might be more difficult, if study and target deviate in structural knowledge about the relevant causal relations (see, e.g., [30] for an analysis of extrapolation in pharmacological research, as well as [26] for a discussion of external validity in the context of risk assessment).

In the interventionist framework, the structural equations at the heart of a causal model support causal explanation from such models as answers to how-questions. It is up to the modeler, of course, to choose a suitable level of abstraction, and to supplement the equations with a meaningful key as to what and how those mathematical functions represent. A causal claim is deemed established if a causal mechanism supports difference-making through intervention. This purely formal, mathematical way of characterizing cause and effect has been criticized by pluralists about causation like Cartwright, who argue that the rich workings of a web of mechanisms should not be reduced away to the model-as-cleanroom idea of computing a function upon intervening on a variable (i.e., setting it to a constant value). Cartwright argues for **thick causal concepts** as an alternative to monolithic, universally applicable, formal and thin accounts of causation: Rich terminology close to experimental practice conveys information in a pragmatically exploitable way:

> [...] there is an untold variety of quantities that can be involved in laws, so too there is an untold variety of causal relations. Nature is rife with very specific causal laws involving these causal relations, laws that we represent

> most immediately using content-rich causal verbs: the pistons *compress* the air in the carburetor chamber, the sun *attracts* the planets, the loss of skill among long-term unemployed workers *discourages* firms from opening new jobs. ... These are genuine facts, but more concrete than those reported in claims that use only the abstract vocabulary of 'cause' and 'prevent'. If we overlook this, we will lose a vast amount of information that we otherwise possess: important, useful information that can help us with crucial questions of design and control. ([4, pp. 814–815])

In embracing rich practice, Cartwright opens up the possibility of importing the connotationally rich lingo of the empirical sciences into the philosophy of science context.

Pharmacology shall serve to illustrate a qualitatively rich frame: Biological (biochemical) networks are a natural way for drug researchers to think about biochemical mechanisms – available mechanistic knowledge is stated in the description of interacting biological pathways. As soon as we enter this level of description, we are dealing with *complex* (non-linear, multi-level) systems that might appear relatively robust w.r.t. higher level functions, but whose inner organization is often considered fragile, input-sensitive, or highly adaptive to changing outer conditions. Moreover, knowledge about such complex systems, if available, often spans multiple levels: Parts of the system might be well-understood at cell level, while others are described in terms of molecular interactions. In addition, some parts in the investigated system lend themselves to being described as inhibitors (with negative influence), others as propagators (passing on positive influence), others again as switches (relaying input in a context-sensitive way) or as catalysts (enabling other parts).[4]

Understanding MECH as built on thick causal concepts, two reasons suggest to keep D-M and MECH apart conceptually: (i) With all the rich terminology introduced above, we can distinguish *difference-making* mechanisms and *non-difference-making* mechanisms (switches under certain circumstances, mechanisms containing cancellation sub-mechanisms, etc.). Importantly, this underlines the distinction between D-M and MECH: Discovering a mechanism might not say anything about D-M if it turns out to be of the non-difference-making sort (see, e.g., [15] for a discussion of mechanistic evidence and masking). (ii) Understanding the existence of a mechanism as *looking into the black-box* and encoding the mechanistic workings of a natural system in a fundamental way, really makes the indicator MECH an entry point for evidence quite different from other information supporting D-M, PROB, and TEMP.

*MECH and PROB.*Saying that a causal mechanism supports a causal relation, entails the possibility that the associated effect is never instantiated if the cause is never triggered, even despite the existence of the supporting mechanism. Such a situation poses difficulties to causal learning techniques, but the **dispositionalist concept of causation** offers an intuitive way of thinking about these situations: A certain drug might have the potential to cause an adverse drug reaction (as shown in controlled tests), but its potential is never unlocked, and its contribution to potential effects is

[4] See, e.g., [36] for an analysis of the behavior of complex systems w.r.t. paradoxical and bidirectional drug effects.

never manifested in circumstances deemed relevant. Explicating a causal association in terms of capacities, propensities, or powers allows for the analysis of causes as dispositional properties (causal tendencies) of a system or of parts of a system to behave in a certain way or develop in a certain direction when a specific trigger manifests. A look at our systematic table shows that the **process theories of causation** (based on mechanisms as processes) swap knowledge about a system's tendency for the counterfactual aspect of D-M. Certain causal questions might lend themselves to a more straightforward treatment in one or the other approach.

And one further point demonstrates the complimentary natures of MECH and PROB: As mentioned above, the **Russo-Williamson Thesis** presents a programmatic call for supplementing difference-making information gleaned from appropriate correlations (tendency and dependency, as encoded in PROB) with mechanistic knowledge (MECH) in data analysis for policy-making. Russo's and Williamson's challenge for decision-makers can be interpreted as saying that complex real-world phenomena can indeed be described by statistics, but can only truly be explained through mechanisms – and responsible decision-making presupposes exactly the possibility of such mechanistic explanation for justification.

*MECH and TEMP.*Interestingly, at least for **thick causal concepts**, knowledge about the mechanistic connection between cause and effect seems to entail knowledge about temporal structure by implicitly conveying information about order, distance, and duration of connected events: (i) The mechanism provides a natural sequential ordering in which a system's components connect and interact, (ii) longer distances never present gaps, but are always filled with mechanistic details (such as a cumulative build-up of a cause component that triggers the effect above a certain threshold), and (iii) when dynamic descriptions of relevant gradients are available, inferential issues due to potentially vague onset and duration of causes and effects vanish (e.g., some pathological condition can be ruled out as an adverse drug reaction, if the condition-specific mechanism suggests the condition's onset before the drug treatment by pointing to a longer incubation time).

As a side remark, it is worth noting that the temporal information conveyed by mechanistic knowledge in the qualitatively rich sense is obviously based on the entrenched compatibility of causal directionality and time's arrow. Although most standard examples of **interventionist accounts of causation** also involve temporally well-ordered events, the mathematically thin, intervention-based explication of cause and effect in such accounts does not prohibit the causal flow to run along *time-reversed* structural equations (causal mechanisms) or to manifest in *simultaneous* causation.

*MECH and the causal claim.*I want to conclude the discussion of the mechanistic connection with remarks on how the relationship between MECH and the causal claim changes with one's choice of causal theory. In the philosophical literature on causal mechanisms, the idea that causal relations are necessarily supported by mechanisms is usually rejected by reference to cases of *causation by omission*: An effect might occur precisely because some potential cause event failed to trigger a certain mechanism ("the flower withered because I did *not* water it"). From a **causal process**

theory perspective, there is prima facie no obvious way to single out a productive mechanism connecting the negative cause and the observed effect, in this case. There is an explication of the situation with emphasis on the counterfactual aspect of D-M, however: The deactivated mechanism can be understood as sustaining some *default* effect state under normal conditions, and the failure to activate this sustaining mechanism now clears the way for other mechanisms that contribute to the effect's deviant state.

Yet again, if one investigates such cases through the lens of an **interventionist causal theory**, matters are different: Since the mechanism for an effect is described as the set of links connecting the cause to its direct parents (i.e., a functional assignment), an omission might be represented by the negative instantiation of a direct cause (i.e., a direct parent variable) – the mechanism is *not deactivated* but *vital* in *determining* the effect's concrete manifestation.

When we turn to medicine one more time, we can see that drawing on **thick causal concepts** helps bridging the gap between negative cause events and their effects: For example, when some neuro-blocker is *not administered*, this might gradually lead to a pathological condition in clearing the way for neurotransmitters to create harmful stress in the body. Such situations are often described in terms of natural gradients or balance/imbalance, and when a drug is well-understood, it is possible to describe the causal chain of events as development over time from the event of *not administering the drug* to the pathological condition. Naturally, all kinds of negative states like *decrease*, *depletion*, *reduction* etc. can play triggering roles in complex interactions of a system's components.[5]

In this section, I specifically contrasted the causal indicator MECH in its thick, domain-specific understanding, encoding biological and physical knowledge, with MECH as a thin concept, encoded in systems of structural equations. Both readings can be made compatible – not within one and the same theory, but in a cross-theoretical view on causal assessment, resting on inference through causal indicators. This not only means that cross-fertilization of seemingly clashing theories is possible (in exchanging complementary information), it might also suggest ways in which domain-specific, context-inducing talk about *compressed air, attracted stellar bodies, and discouraged employers* may eventually be systematized a bit further to allow for knowledge transfer between contextual frames.

[5] Moreover, in the context of pharmacological risk assessment one is usually presented with positive input: Our starting point is the suspicion of a causal connection between a substance and a side-effect – what does the mechanism between the drug and its effect look like? The answer might have to do with balance and imbalance, too, especially if the drug acts as an inhibitor. But this inhibition can again be understood as contributing to a complex network of causal interactions on the bio-chemical level. So, for the purpose of relating the causal claim and the MECH indicator in the context of pharmacological risk assessment, hope certainly is, that for a causally efficacious substance, scientists will *eventually* find a causal mechanism as a satisfactory answer to the "How?". This is by no means guaranteed, however: It is well known, for example, that putting newborns on their back prevents fatal incidents subsumed under the *Sudden Infant Death Syndrome (SIDS)* – yet a (mechanistic) explanation is still lacking after many years of research.

2.4 Causes precede their effects (TEMP)

*TEMP and the causal claim.*Since temporal precedence as an indicator of causation seems like the weakest indicator of the four, my comparative listing concludes with the question, how TEMP might fail in supporting the causal claim when indeed it holds. In other words: Under what circumstances might TEMP falsely mistake an effect for the cause? Three cases shall be discussed briefly: (i) instances of backward causation, (ii) feedback loops, and (iii) falsely determined event durations:

(i) Backward causation, a philosophical problem: This case is just another reminder of the ramifications of theory choice in causal assessment. In principle, if the theory permits effects that precede their causes, then the indicator must be adjusted accordingly, and the framework tweaked to be sensitive to signals of backward causation. Nevertheless, to be fair: This case is repudiated by the majority of philosophers.

(ii) Feedback loops, a theoretical problem: If TEMP receives information of mismatching granularity, i.e., information about generic variables on a coarser level than the actual investigation, inferred patterns might show loop structures in which cause and effect mutually influence each other, thus reversing the cause–effect relation.

(iii) Event durations, a practical problem: In defining or determining event boundaries in a given study, cause and effect might be mistaken when the onset of the actual cause is erroneously located after the actual effect event. E.g., in medicine, a disease might have a longer incubation time (i.e., an earlier onset) than recognized, but its early, undetected symptoms might already have effects which might then mistakenly be deemed causes of the disease's outbreak (based on covariation).

All these cases once more emphasize the virtues of basing one's causal assessment on multiple, distinct indicators which mediate between the hypothesized causal association and incoming evidence and do so by marking the theoretical demands of the chosen causal theory.

3 Summary and conclusions

In my discussion of the ramifications of theory choice in causal assessment, I first collected the pivotal causal indicators found in the prominent theoretical approaches towards the concepts of cause and effect. In doing so, I distilled three main goals of the hunt for causal knowledge: control, prediction, and explanation. Formally, these goals surface as control and counterfactual reasoning (difference-making formalisms), dependency/tendency (probabilistic terminology), and mechanistic knowledge (explicated as systems and their properties). As a fundamental further category I added temporal structure – a dimension more or less explicitly debated throughout the literature. In order to relate these four aspects to one another and to the hypothesized causal association, I expressed them as *indicators of causation* (of different strength) at a level of abstraction suitable for the purpose of this paper. Delineating the indicators in such a way allowed me to discuss their interdependencies through the lenses

of the most prominent theories of causation, listed in Tab. 1. The table is certainly not exhaustive, but it illustrates how the theoretical differences between prominent analytic theories of causation can be broken down in terms of the four indicators. Let me reiterate that, notably, all of the theories discussed rely on a combination of two aspects (explicitly or implicitly).

Putting on different 'theoretical hats', I revisited D-M (Subsec. 2.1), PROB (Subsec. 2.2), MECH (Subsec. 2.3), and TEMP (Subsec. 2.4) in order to compare the ramifications of choosing contrasting theoretical viewpoints. Naturally, theories simplify – and as my tour through the theoretical landscape evinced, each of the theoretical stances on its own has its deficits. Uncounted philosophical disputes in the vast literature on the concept of causation testify this finding. Yet, understanding the fundamental aspects encoded in D-M, PROB, MECH, and TEMP as *genuine indicators of causation* facilitates a unifying view on different theories each featuring different strong points. Different theoretical/philosophical approaches might not be able to jointly enter in the direct formulation of causal claims – different approaches come with different languages, as demonstrated. Different approaches might enter in the formation of causal indicators, though, and contribute to the assessment of a hypothesized causal association on this higher level (given they are semantically compatible).

As Tab. 1 makes visually graspable, no theory (amongst the ones discussed) makes use of all four indicators. Since I derived D-M, PROB, and MECH from instruction, prediction, and explanation as main targets of causal reasoning, my discussion shows that (and in which way) any given theory might be good at two of these main targets, but does not exhaust the full spectrum. But of course, scientific inquiry in all its richness builds on heterogeneous input (across all four indicators) on the one hand and aims at meaningful, actionable output (across all main targets) on the other hand: Machine learning algorithms trawling unstructured bodies of data (PROB) must be supplemented with heuristics and expert knowledge (MECH) to generate actual insights; in the clinical context, biomedical knowledge about a system (MECH) only enters in causal inference if the interaction of the system's components make a difference (D-M) to the patient; and for practical policy-making, instructions as to how to intervene in something as complex as society, the potential difference-making power of an intervention (D-M) only brings to bear if a policy with the right independencies (PROB) is achievable. Adopting the unificatory, cross-theoretical view I advance here enables the researcher to draw on heterogeneous resources and paves the way for a novel, conciliatory perspective on seemingly contradictory theoretical stances: Just imagine a *small miracle* and picture yourself in the peaceful world where Pearl tells you how to control your effects, Good calculates the causal tendency, and Cartwright provides a thick explanation. This world is not far.

References

1. Anjum, R.L., Mumford, S.: Causal dispositionalism. In: A. Bird, B. Ellis, H. Sankey (eds.) Properties, Powers and Structure, chap. 7, pp. 101–118. Routledge (2012)
2. Beebee, H., Hitchcock, C., Menzies, P. (eds.): The Oxford Handbook of Causation (Oxford Handbooks) (2009). URL http://amazon.com/o/ASIN/019927973X/
3. Cartwright, N.: What is wrong with bayes nets? Monist **84**(2), 242 (2001)
4. Cartwright, N.: Causation: One word, many things. Philosophy of Science **71**(5), 805–819 (2004). URL http://www.jstor.org/stable/3693306
5. Cartwright, N.: Are RCTs the Gold Standard? Biosocieties (2), 11–20 (2007). URL http://dx.doi.org/10.1017/S1745855207005029
6. Cartwright, N.: Causal Powers: What Are They? Why Do We Need Them? What Can be Done with Them and What Cannot? Tech. Rep. 04/07 (2007)
7. Cartwright, N., Stegenga, J.: A Theory of Evidence for Evidence-Based Policy. In: P. Dawid, M. Twinning William Vasilaki (eds.) Evidence, Inference and Enquiry, chap. 11, pp. 291–322. OUP (2011)
8. Dowe, P.: Causal Process Theories, chap. 10, pp. 213–233. In: Beebee et al. [2] (2009). URL http://amazon.com/o/ASIN/019927973X/
9. Dummett, A.E., Flew, A.: Symposium: Can an effect precede its cause? Aristotelian Society Supplementary Volume **28**(1), 27–62 (1954)
10. Glennan, S.: Mechanisms, chap. 15, pp. 315–325. In: Beebee et al. [2] (2009). URL http://amazon.com/o/ASIN/019927973X/
11. Glennan, S.: Mechanisms, causes, and the layered model of the world. Philosophy and Phenomenological research **81**(2), 362–381 (2010). URL http://dx.doi.org/10.1111/j.1933-1592.2010.00375.x
12. Good, I.J.: A theory of causality. British Journal for the Philosophy of Science **9**(36), 307–310 (1958)
13. Halpern, J.Y., Pearl, J.: Causes and explanations: A structural-model approach. part i: Causes **56**(4), 843–887 (2005). URL http://www.jstor.org/stable/3541870
14. Hume, D.: An Enquiry Concerning Human Understanding. The University of Adelaide Library 2004 (derived from the Harvard Classics Volume 37, 1910 P.F. Collier & Son.) (1748). URL http://ebooks.adelaide.edu.au/h/hume/david/h92e/
15. Illari, P.M.: Mechanistic evidence: Disambiguating the russo–williamson thesis. International Studies in the Philosophy of Science **25**(2), 139–157 (2011)
16. Kerry, R., Eriksen, T.E., Lie, S.A.N., Mumford, S.D., Anjum, R.L.: Causation and evidence-based practice: an ontological review. Journal of Evaluation in Clinical Practice **18**(5), 1006–1012 (2012). URL http://dx.doi.org/10.1111/j.1365-2753.2012.01908.x
17. Kitcher, P.: Explanatory unification and the causal structure of the world. In: P. Kitcher, W. Salmon (eds.) Scientific Explanation, pp. 410–505. Minneapolis: University of Minnesota Press (1989)
18. Landes, J., Osimani, B., Poellinger, R.: Epistemology of Causal Inference in Pharmacology: Towards a Framework for the Assessment of Harms. European Journal for Philosophy of Science (2017). DOI 10.1007/s13194-017-0169-1
19. Lewis, D.: Causation **70**(17), 556–567 (1973)
20. Lewis, D.: Causation as influence **97**(4), 182–197 (2000). URL http://www.jstor.org/stable/2678389
21. Machamer, P., Darden, L., Craver, C.F.: Thinking about Mechanisms. Philosophy of Science **67**(1), 1–25 (2000). URL http://www.jstor.org/stable/188611
22. Mackie, J.L.: Causes and conditions. American Philosophical Quarterly **2**(4), 245–264 (1965). URL http://www.jstor.org/stable/20009173
23. Mumford, S.: Causal Powers and Capacities, chap. 12, pp. 265–278. In: Beebee et al. [2] (2009). URL http://amazon.com/o/ASIN/019927973X/
24. Mumford, S., Anjum, R.L.: Getting Causes From Powers. Oxford University Press (2011)
25. Osimani, B.: Hunting side effects and explaining them: Should we reverse evidence hierarchies upside down? Topoi **33**(2), 295–312 (2014)
26. Osimani, B., Mignini, F.: Causal assessment of pharmaceutical treatments: Why standards of evidence should not be the same for benefits and harms? Drug Safety (1), 1–11 (2015). DOI 10.1007/s40264-014-0249-5
27. Pearl, J.: Causal diagrams for empirical research. Biometrika **82**(4), 669–688 (1995)
28. Pearl, J.: Causality: Models, Reasoning, and Inference, 1 edn. Cambridge University Press (2000)
29. Pearl, J.: Causality: Models, Reasoning, and Inference, 2 edn. (2009)

30. Poellinger, R.: Analogy-Based Inference Patterns in Pharmacological Research (2017). Forthcoming
31. Popper, K.R.: The propensity interpretation of probability. The British Journal for the Philosophy of Science **10**(37), 25–42 (1959). URL http://www.jstor.org/stable/685773
32. Price, H.: Time's Arrow & Archimedes' Point: New Directions for the Physics of Time. Oxford University Press (1996)
33. Reichenbach, H.: The Direction of Time. University of Los Angeles Press (1956). URL http://plato.stanford.edu/entries/physics-Rpcc/
34. Russo, F., Williamson, J.: Interpreting Causality in the Health Sciences. International Studies in the Philosophy of Science **21**(2), 157–170 (2007). DOI 10.1080/02698590701498084
35. Salmon, W.: Scientific Explanation and the Causal Structure of the World. Princeton University Press (1984)
36. Smith, S.W., Hauben, M., Aronson, J.K.: Paradoxical and bidirectional drug effects. Drug Safety **35**(3), 173–189 (2012). DOI 10.2165/11597710-000000000-00000. URL https://doi.org/10.2165/11597710-000000000-00000
37. Spirtes, P., Glymour, C., Scheines, R.: Causation, Prediction, and Search. Adaptive Computation and Machine Learning. MIT Press (2000)
38. Spohn, W.: Causation: An alternative **57**(1), 93–119 (2006). URL http://www.jstor.org/stable/3541654
39. Stegenga, J.: Down with the Hierarchies. Topoi **33**(2), 313–322 (2014). URL http://dx.doi.org/10.1007/s11245-013-9189-4
40. Strevens, M.: Depth: An Account of Scientific Explanation. Harvard University Press (2008)
41. Suppes, P. (ed.): A Probabilistic Theory of causality. North-Holland Pub. Co. (1970)
42. Verma, T., Pearl, J.: Causal networks: Semantics and expressiveness. In: Proceedings of the 4th Annual Conference on Uncertainty in Artificial Intelligence (UAI-88). Elsevier Science, New York (1988)
43. Williamson, J.: Probabilistic Theories, chap. 9, pp. 185–212. In: Beebee et al. [2] (2009). URL http://amazon.com/o/ASIN/019927973X/
44. Woodward, J.: Making Things Happen: A Theory of Causal Explanation (Oxford Studies in the Philosophy of Science). Oxford University Press (2003)
45. Worrall, J.: Why There's no Cause to Randomize. British Journal for the Philosophy of Science **58**(3), 451–88 (2007). URL http://dx.doi.org/10.1093/bjps/axm024