The Worldly Infrastructure of Causation

Naftali Weinberger, Porter Williams, and Jim Woodward

May 22, 2023

Abstract

This paper describes an alternative to currently dominant philosophical approaches to the metaphysics of causation. It is motivated by the gap that currently exists between metaphysical accounts and recent epistemological research on causal reasoning and methods for discovering causal relationships. Our approach aims at characterizing structural features of the actual world that support, and are exploited by, successful strategies for causal reasoning and discovery. We call these features the "worldly infrastructure" of causation. We identify several elements of this worldly infrastructure, sketch an account of their physical bases, and explain how they contribute to the possibility of successful causal reasoning.

1 Introduction

Recent work on causation has taken a variety of forms. Researchers in statistics, econometrics, and machine learning have been mainly interested in epistemological and methodological issues surrounding causal inference – issues concerning how one can reliably infer causal conclusions from various sorts of data. Examples include the constraint-based causal discovery methods in (Spirtes et al., 2000), results about identifying causal effects (e.g., (Pearl, 2009)), machine learning techniques for inferring causal direction (Peters et al., 2017) and the potential outcome frameworks employed by many social scientists (Rubin, 1974; Hernán and Robins, 2020). Many philosophers – especially those focusing on causal explanation – have devoted significant attention to these methods. But many have instead been concerned with the metaphysics of causation. This work also takes a number of forms. Some hold that the metaphysics of causation requires the introduction of special entities – powers, capacities and the like. Others reduce causal claims to counterfactuals and elucidate the latter through possible worlds semantics. Still others propose that causal claims be understood via their relation to laws of nature. In addition, some philosophers claim that causation or causal relationships can be identified with processes or relationships in fundamental physical theories so that, metaphysically, causation is just, e.g., transfer of energy and momentum. Common to all of these efforts is a search for "truth conditions" or "grounds" for causal claims and/or attempts to specify what causation "is" or what in the world "corresponds" to the causal nexus.[1]

These metaphysical projects are conducted with little or no connection to the work on the epistemology of causation referenced above. Many metaphysicians do not acknowledge this work, much less attempt to integrate it with their proposals. Indeed, many contemporary metaphysicians insist that analyzing the metaphysics of

[1]Still other authors propose to analyze "our concept" of causation, typically with the aim of reducing that concept to some non-causal ingredients. We will not discuss these projects except to say that the infrastructure project we describe is also distinct from them. In particular, we do not consider the infrastructure features we discuss to be part of "our concept" of causation or the meaning of the word "cause".

causation can, and perhaps should, be sharply separated from the study of epistemological strategies for discovering causal relationships – the latter regarded as being of "merely" methodological or practical interest.

A recent prize-winning book (Paul and Hall, 2013) is an exemplar of this general practice. They aim to reduce garden-variety causal claims to counterfactuals, with their semantics determined by choosing a Cauchy surface on which the antecedent is true and evolving it forward in time using equations of motion of fundamental physical theories (following Maudlin (2007, chapter 1)). Ordinary subjects making causal judgments obviously don't do so by anything like this procedure, but Paul and Hall provide no epistemological story about how, if this is what causal claims are, people are able to reason to correct causal judgments.[2] There is no effort to make a connection with the large empirical literature on how scientists or laypeople draw causal conclusions. We single out Paul and Hall not for special criticism, but because their book is a well-known and particularly thoughtful representative of the general practice of analyzing the metaphysics of causation independently of epistemological work on how causal claims are established.[3]

Our goal in this paper is to describe and develop a project that is different from the metaphysical projects described above, but is also not properly described as epistemology. It represents a third possibility. This project aims to elucidate what we call the "worldly infrastructure" underlying the application of causal concepts and

[2] Paul and Hall do discuss, sometimes approvingly, the use of resources from causal modeling to represent and clarify causal claims, but not for causal discovery. It is precisely those epistemological strategies for learning about causal relations that are central in the causal modeling literature but play little role in mainstream work on the metaphysics of causation.

[3] This separation is facilitated by Paul and Hall's focus on "actual cause" claims, for which there is no consensus on an appropriate discovery methodology. The epistemological strategies and infrastructure features to which we draw attention are primarily exploited to identify type-level causal relations.

causal reasoning. The basic idea is this: there are certain generic features of our world that license and support the application of causal thinking and inferences to causal conclusions. These features include (but are not limited to) the following: (i) some variables are statistically independent of others (not everything is correlated with everything else); (ii) interventions, in the sense of unconfounded manipulations, are often possible and, more generally, many systems exhibit naturally occurring exogenous sources of variation that count as interventions in the technical sense even if they do not involve human manipulation; (iii) the macroscopic, coarse-grained behavior of many systems is largely independent of variations in their microscopic realizing details and this allows for "unambiguous" interventions into such systems and the discovery of robust causal generalizations about their macroscopic behavior.

A motivating assumption of our project is that our concepts and strategies for causal reasoning developed to exploit the fact that we live in a world in which these generic features obtain. For example, the truth of (ii) is one factor that contributed to our developing a notion of causation that is closely linked to what happens under interventions – (ii) helps to ensure that this intervention-based notion will be useful. If interventions were rarely or never possible, we would presumably not have developed a notion of causation tied to interventions. This is the sort of thing we have in mind when talking about the worldly infrastructure that supports causal thinking.

Our project is also motivated by the empirical success of the causal discovery methods mentioned above. These methods reliably deliver truths about causal relations and it is natural to ask why they are successful. Our answer is that these causal discovery methods work because they rely upon, and exploit, the worldly infrastructure we describe. Insisting that these methods are of no "metaphysical" significance denies any tight connection between the success of these methods and the structure of the world, a position we consider untenable. We consider it a virtue that our project is responsive to the epistemology and methodology of causal inference and identifies features of the world that explain the success of those methods; our discussion of the "worldly infrastructure" catalogues some of these features. As we will argue, the claim that the success of causal discovery methods supports the structural presuppositions underlying those methods has much in common with similar inferences commonly made elsewhere in science.

The project of elucidating the worldly infrastructure supporting causal reasoning differs in key respects from familiar metaphysical projects. It does not require special metaphysical concepts (e.g., powers) but employs only notions already in the toolbox of the practicing scientist, such as statistical independence and exogeneity. It does not proceed by appealing to intuitions about cases. It aims, first and foremost, at characterizing the contingent structure of the actual world that supports causal reasoning. Accordingly, we make no claim that the causal concepts employed in our world are metaphysically necessary; whether they apply in other worlds – or even to all phenomena in the actual world – depends on whether the worldly infrastructure obtains. We offer more extended discussion of these and other differences in section 6.

These differences may suggest that we are not providing "genuinely" metaphysical information about causation. We have little interest in a dispute about what counts as genuine metaphysics. What is important is that the infrastructure project characterizes objective structural features of the world on which causal reasoning relies. Exploring the worldly infrastructure of causal reasoning, while different from projects pursued within mainstream metaphysics of causation, is of considerable importance in its own right.

2 The General Framework

The framework we adopt has several parts. First, there are the infrastructure features themselves, such as (i)-(iii) above. Second, there are connecting principles that license inferences from the presence of some particular infrastructure, typically in conjunction with other information, to causal conclusions. In this sense the connecting principles reflect the use of the infrastructure in causal inference. The general form of the connecting principles is:

> **P**: (a) Particular instances of infrastructure features obtain (e.g., variables exhibiting some pattern of statistical independence and dependence, some manipulation of X with respect to Y has the characteristics of an intervention, etc.) + (b) additional information (possibly causal in character) $\longrightarrow$ causal conclusion.

For example, the claim (1) that the world is often such that interventions are possible is a general claim about an infrastructure feature. The claim (2) that some particular manipulation of X with respect to Y satisfies the conditions for an intervention is a claim of form (a). The characteristic interventionist claim (**M**) – if Y systematically changes under some intervention on X, then X causes Y – is a connecting principle of form **P**. This licenses a causal conclusion if the antecedent of this conditional is satisfied. For (**M**), the additional information (b) referred to in **P** is built into the characterization of an intervention, which includes causal information about the effect of an intervention I on X.

Perhaps (**M**) could be true even if (1) is false, but if (1) were false then (**M**) would be largely useless as a principle in causal inference since its antecedent will rarely be satisfied. Assuming that our ways of thinking about causation have developed because they are functional or useful in some way (Woodward, 2014), it is hard to see why we would have developed a notion of causation in which (**M**) plays a central role if (1) were false. In reality (1) is true, so it has been useful for us to develop a notion of causation that is intimately connected with what happens under interventions. This is one illustration of what we have in mind when we say that our thinking about causation is formed in response to, and exploits, the presence of the infrastructure features.

It may be helpful to clarify the functional approach to causal reasoning with an analogy. When disambiguating visual scenes, the human visual system relies on the built-in assumption that illumination is generally from above (e.g., from the sun) and that objects tend to be convex. That this assumption is generally correct partially explains why the visual system is generally reliable when inferring shape from shading. (It also partially explains why the human visual system developed to exploit it.) We see the existence of the causal infrastructure analogously: the worldly infrastructure itself is analogous to the physical facts about illumination from above and the convexity of objects, while the inferential principles that exploit that infrastructure to deliver causal conclusions are analogous to the visual system that exploits those physical facts to deliver reliable inferences of shape from shading.

The infrastructure features are related to each other in a way that supports consistent causal reasoning and inference. That is, we can exploit the presence of distinct infrastructure features in ways that converge on consistent causal conclusions – conclusions that thus can be transferred across different inferential contexts. As an illustration, consider the principle of the common cause (CC): if X and Y are statistically dependent, then either X causes Y, Y causes X, or X and Y have a (set of) common cause(s). Perhaps one could imagine a world in which CC regularly fails when "cause" is understood along interventionist lines: a world in which frequently X and Y are statistically dependent but interventions on X are not associated with changes in Y, interventions on Y are not associated with changes in X, and there is no third variable Z such that interventions on Z are associated with changes in X and Y. Whether or not such a world is possible, our world is not like this. In our world, there are systematic connections between patterns of statistical association and what will happen under interventions that are captured by the principle of the common cause – if X and Y are statistically dependent then one (or more) of the three possibilities allowed by CC follows. This means that in our world, there is a connection between what may be inferred from statistical dependencies and what will happen under interventions. Given CC and that X and Y are statistically dependent, if we determine that X and Y do not have a common cause and Y does not cause X, we may infer that some intervention on X is associated with Y. (Here we employ a connecting principle of form **P** which uses CC, information about statistical dependence, and information about the absence of causal relations to infer to a causal conclusion.) Again, our ways of thinking about causation have developed to take advantage of such connections.

As another illustration, consider the Causal Markov Condition (CMC). The CMC is satisfied if, for every variable X_i in a causal graph $\mathcal{G}$, conditioning on the parents (direct causes) of X_i renders X_i statistically independent of every other variable in $\mathcal{G}$, except possibly its descendants (effects). It is crucial for usefully applying CMC that statistical dependencies – unconditional and conditional – not be ubiquitous. Suppose instead that for any set of variables $\mathbf{X} = \{X_1, \ldots, X_n\}$ we could measure, each X_i and X_j were statistically dependent and remained statistically dependent as we conditioned on all subsets of other variables in the set $\mathbf{X}$. All fully connected graphs (every variable directly causally connected to every other) are consistent with CMC, so CMC alone would be of little use. Normally we invoke additional conditions that justify choosing sparser graphs over fully connected ones. For example, the widely used faithfulness condition selects only graphs that entail, solely from their structure, all the conditional and unconditional independence relationships in the probability distribution. However, if for every set of variables $\mathbf{X}$, no such independence relationships obtain, conditions like faithfulness even in conjunction with CMC would not be useful for inferring anything definite about causal structure. Fortunately, in our world there is a considerable amount of unconditional and conditional statistical independence and we regularly exploit such independence, in conjunction with other conditions (like CMC and faithfulness), to learn about causal relations.

The existence of variables that are statistically independent is also required for other familiar causal inference procedures like the use of randomized experiments – we cannot randomize if there are no natural statistical independencies among the variables of interest and we cannot produce any. There would be no possibility of causal learning on the basis of such experiments. Again, we should distinguish between (i) the connecting principle that licenses causal conclusions from the results of randomized experiments: roughly, if in an experiment in which assignment of C is randomized, the incidence of E is higher in the treatment group than in the control group (and we rule out that this a statistical fluke), then infer that C causes E; and (ii) the infrastructure feature: statistical independencies that enable randomization are generically available. That the infrastructure (ii) is required for the strategy (i) to be used illustrates the general fact that to profitably apply our causal reasoning strategies, the worldly infrastructure supporting those strategies must be present.

3 Some examples of worldly infrastructure

We now turn to some candidates for the worldly infrastructure of causation (including some already introduced). To avoid pedantry, we will use "variable" to describe both elements of the world standing in causal relations and representations of those elements.[4]

3.1 Statistical Independence

The world contains many pairs (triples, n-tuples) of variables that are strictly statistically or probabilistically independent of each other, and many others that "effectively independent," i.e., sufficiently close to independence that they can be treated as independent for many inferential purposes. Some examples: (A) The color of our respective socks on any given day is independent of whether some randomly selected Parisian had eggs for breakfast that day. (B) The outcomes of successive coin tosses with the same generating set up are typically effectively independent. Those outcomes are also independent of many other variables: the time of day at which the tosses occur, fluctuations in stock prices, etc. (C) Mendel's law of independent assortment states that alleles for different traits are passed to offspring independently of each other. This "law" does not always hold (because of genetic linkage, among other considerations) but when it does, the independence relations it generates can be exploited in causal inference, as in the use of "Mendelian Randomization" to make inferences about the role of environmental exposures in causing diseases (Lawlor et al., 2008; Sanderson et al., 2022). (D) Assumptions that the velocities of any two molecules in

[4]The reader is free to use "quantity" (where quantities can be two-valued, real-valued, etc.) when the discussion concerns causal relata in the world and reserve "variable" for representations of quantities.

a dilute gas are statistically independent immediately prior to their collision – like the Stosszahlansatz or the assumption of "molecular chaos" – have played an important role in statistical physics since the time of Boltzmann (Brown et al., 2009).

We have already described several ways in which the presence of statistical independence is exploited in causal inference. As additional illustration, the standard proof of CMC assumes that each variable is a deterministic function of other measured variables (its parents) and an additive "error" term. The error term for each equation is assumed to be probabilistically independent of the parents and the error terms across equations are assumed to be independent of each other.

Similar independence assumptions concerning error terms are often made when causal modeling or structural equations frameworks are employed to identify causal relations, even if CMC is not explicitly used. For instance, applications of the "back-door criterion" of Pearl (2009, p. 79) identify the effect of X on Y by conditioning on variables along any confounding paths linking X and Y in a way that ensures any probabilistic dependence between the variables after conditioning reflects their causal relationship. This requires that absent the confounding path and the causal relationship, X and Y would be probabilistically independent. Identifiability conditions like the back-door criterion are crucial in the social sciences, since they are required for specifying which variable sets suffice to control for confounding. If causal independence did not typically produce probabilistic independence, standard techniques for identifying and estimating causal relations could not be reliably applied.

As yet another illustration, techniques for inferring causal direction often exploit information about statistical independence. As a simple example, consider a system of three variables, with X independent of Y but X and Z dependent and Y and

Z dependent. Then a fairly reliable heuristic delivers the judgment that X and Y cause Z: the causal direction goes from X and from Y to Z, with X and Y causally independent.[5] Obviously we could not develop or use this heuristic if all or most variables were pairwise statistically dependent.

Looking beyond explicit frameworks for causal discovery, the factorizability of joint probability distributions (reflecting statistical independence) is commonly understood as indicating causal independence in science. For instance, the factorization of a joint distribution in physics is commonly understood to reflect facts about (effective) causal independence of the subsystems (i.e., the absence of nontrivial physical influence between the subsystems). Such inferences are ubiquitous in classical physics and, although quantum entanglement makes the connection subtle, remain extremely common in quantum physics as well. This has important methodological consequences, not least that knowing that a joint probability distribution factorizes often justifies analyzing the causally independent subsystems independently, greatly facilitating computation. For example, quantum field theories are required to satisfy the cluster decomposition principle: roughly speaking, the principle requires any two events occurring at sufficiently large spatial separation to be probabilistically independent (Weinberg, 1995, chapter 4). This is meant to capture the fact that experiments at Fermilab are causally independent (and thus probabilistically independent) of anything taking place in the accelerator tunnel at SLAC. It speaks to the ubiquity of this type of reasoning in physics that one of the core principles of quantum field theory is essentially a formalization of this relationship between causal independence and the factorization of joint probability distributions.[6]

[5] See (Woodward, 2022a) for more on why this and other, more sophisticated procedures for inferring causal direction work.

[6] A completely satisfactory characterization of the physical significance of cluster decomposition is surprisingly subtle; see (Dougherty et al., 2023).

3.2 The possibility of interventions

It is frequently possible to intervene on some variables with respect to others, or to find naturally occurring variables that have the characteristics of interventions. (By "intervention" we mean an unconfounded exogenous manipulation that satisfies criteria like those described in (Woodward, 2005).) Built into the possibility of an intervention is that it is possible to intervene on some variables in a way that is independent (causally and statistically) of the values of other variables, and in a way that influences certain other variables, if at all, only indirectly.[7] The world is such that it does not conspire to make such interventions impossible. For example, the world is not superdeterministic: that would entail that whenever a researcher does an experiment in which she thinks she is performing an intervention I that manipulates X, and Y changes, this is actually due to some unobserved common cause of I and Y, or due to some common cause of X and Y that "just happens" to be correlated with I. The upshot is that the supposed intervention I is not really an intervention at all and the conclusions the researcher draws about the causal relation between X and Y are mistaken. The assumption that this sort of systematic confounding is not widespread is required if we are to draw correct causal conclusions from experiments. Conversely, it might be argued that since we apparently can reason correctly from experimental results to causal conclusions, this puts pressure on the possibility of superdeterminism – to accommodate this fact, superdeterminism would have to incorporate some very remarkable kinds of pre-established harmony to generate this appearance in the absence of causal connections.[8]

[7]It is also generally assumed that the world is such that *local* interventions are possible. This means, roughly, that when it is possible to intervene with I on X, I itself does not consist of some delicately tuned manipulation of many other, perhaps spatially separated, variables.

[8]Superdeterminism is different from – and much stronger than – the possibility that the world is governed by deterministic laws. The latter poses no particular problems for causal inference.

Going further, often variables occur "naturally" that have the characteristics of interventions (i.e., that are "exogenous" to some candidate causal relationship from X to Y and that provide sources of variation in X that are associated with variation in Y only through the variation in X they produce). Such variables greatly facilitate causal inference, making possible "natural experiments" in many areas of science and the identification of variables that can serve as "instruments" (i.e., instrumental variables) for investigating causal relationships (Angrist et al., 1996). These techniques illustrate, among other things, that even if the relationship between X and Y is confounded (by a common cause Z, for example), as long as we can find an instrumental variable W that is an additional source of variation in X, where this variation is independent of the variation in X caused by Z, we can use this information to detect whether there is a causal relation between X and Y. Variables having these intervention/instrument-like characteristics are not always available, but often they are (even if identifying them requires considerable ingenuity) and can be exploited in causal inference. As with other infrastructure features, the existence of such variables is a broadly empirical fact about our world.

A closely related observation is that it is sometime possible to "fix" the values of variables exogenously by imposing values that remain unchanged, via processes that are themselves causally uninfluenced by other variables in the system. Consider the ideal gas law, (IG) $PV = kT$, which relates the pressure P, volume V, and temperature T of a gas at thermodynamic equilibrium. (IG) is silent about the causal relationships among these variables. Now suppose the gas is confined to a box of fixed volume and immersed in a heat bath of fixed temperature. The heat bath fixes the value of T exogenously (since the other variables V and P do not influence the heat bath), while the volume of the gas is fixed exogenously by enclosing it in a rigid container. We can represent this with the equations (2) $V = v$ and (3) $T = t$ indicating that these variables have been set to fixed values. We can then use (2) and (3) to solve for the pressure using (IG). In a more general setting, Simon (1953) used such facts about the order in which the values of variables in a set of equations can be determined to infer the causal ordering (i.e., which variables cause others). In the specific experimental setup we consider, this licenses the inference that V and T are causes of P.

3.3 Connections between Causal and Statistical Relations

We noted above that in our world there are systematic connections between statistical independence and dependence relations (both conditional and unconditional) and causal relationships, as reflected in principles like CC and CMC. These principles play an important role in learning about causal relationships. Indeed, without something like these principles it is hard to see how one could learn about causal relationships from observation of statistical relationships, as we sometimes clearly can. However, it is difficult to argue that they reflect metaphysically necessary features of causation; even those who aim to identify metaphysically necessary features of causation generally don't claim that CC or CMC are among those features. We think that the moral to draw from this is not that CC or CMC play no important role in understanding causation, but that this narrow conception of what qualifies as metaphysically important information about causation omits important aspects of how causal reasoning relates to features of the actual world. Our suggestion that CC and CMC reflect features of the worldly infrastructure that support the application of causal reasoning is meant to capture this.

The general usefulness of CC and CMC is not threatened by the possibility that they are not satisfied by certain phenomena (such as measurement results on entangled quantum systems).[9] In line with our characterization of these principles as reflecting contingent facts about the infrastructure, however, we do not view the mere possibility that one could imagine cases in which they are violated as a genuine limitation.[10]

3.4 Modularity

Say that a representation of a system represents it as modular if certain causal relationships in the system will remain stable or unchanged under modifications of other causal relationships represented in the system. (Modularity is not a binary property but comes in degrees.) For example, a system of two coupled springs is modular to the degree it is possible to modify the relationship $\mathbf{F} = -k_1 x_1(t)$ governing one spring (e.g., by stretching) while leaving the relationship characterizing the other spring unchanged. Say that a system is modular if it has a modular representation, at some level of description, that correctly predicts the results of interventions on variables in the system. It is implausible that it is somehow metaphysically necessary that there exist modular representations for all systems of interest and nothing guarantees that such representations will exist at any chosen level of description. It may well be that some systems are fundamentally non-modular, i.e., they have no predictively accurate

[9]There is an enormous literature, and no consensus, about how Bell-type correlations bear on principles of causal inference. For instance, Glymour (2006) thinks they violate CMC, Wood and Spekkens (2015) think they violate faithfulness-type conditions, and Hausman and Woodward (1999) deny that interventions of the required type are well-defined for entangled systems, so a causal analysis of Bell-type correlations fails even to get off the ground.

[10]For example, Cartwright (1999) imagines a hypothetical chemical factory where they are violated and Elliott and Lange (2022) imagine a world containing only two particles that are (by construction) causally independent but correlated in their movements, but it is not clear that there are realistic cases (at least in the macroscopic world).

modular representation at any level of description. Nevertheless, we are sometimes able to discover modular representations of systems and this greatly facilitates causal analysis. For example, if I know that intervening to knock out one gene in a genetic regulatory network will not lead the entire system to reorganize so that causal relations throughout the network change (which would be a massive failure of modularity), this makes it much easier to learn about the causal structure of the network. (Many investigations of genetic regulatory networks assume this sort of modularity, often with good empirical support.)[11] Again, it is a fact about our world that many systems of causal relations are modular at some accessible level of measurement and description. This is another feature of the worldly infrastructure supporting causal inference.

3.5 Value-Relation Independence

Our discussion of principles such as CC and CMC has emphasized "constraint-based" methods for learning causal structure based on conditional independence relations, but there now exist machine learning methods for causal inference that exploit properties of a probability distribution beyond factorization. Central to these methods is the "principle of independent mechanisms" (PIM) (Peters et al., 2017, chapter 2.1). Consider a system of two random variables in which X causes Y; PIM formalizes

[11] Again, it may well be that some genetic systems that are non-modular, or at least are best represented as non-modular at levels of description appropriate for genetics. Mitchell (2009) (following Greenspan (2001)) describes a hypothetical gene network in which interventions on one gene or node produces a change in causal relations throughout the network. We agree that there may be real gene networks that behave this way. Insofar as it makes sense to talk about "the causal structure" of such networks, our point is that learning this structure is going to be much more difficult than learning the structure of a modular network. Moreover, if the failure of modularity is too massive – e.g., the network completely reorganizes in different ways depending on which node is intervened on – it becomes an open question whether it is useful or appropriate to talk about the causal structure of the network. Alternatively, this sort of massive failure of modularity might be taken to signal that we are analyzing the system at the wrong level of description.

the sensible expectation that the "mechanism" that determines the probability distribution over the cause variable X operates independently of the "mechanism" that determines the conditional probability distribution over the effect variable Y, given X. More concisely, PIM requires that $\Pr(Y \mid X)$ be "independent" of $\Pr(X)$. Here "independence" clearly cannot mean statistical independence.[12] Rather, it means something like $\Pr(Y \mid X)$ and $\Pr(X)$ can vary independently of one another, in the sense that $\Pr(Y \mid X)$ will not change under suitable changes in $\Pr(X)$ (changes in $\Pr(X)$ produced by interventions, roughly speaking), and vice versa. This can be understood as a modularity condition (since it asserts that the mechanism generating X can be modified independently of the mechanism connecting X to Y) and reflects an assumption found throughout science that laws and initial conditions are similarly modular.[13]

One virtue of PIM is that by exploiting information about probability distributions that goes beyond that used by constraint-based causal discovery methods, it can solve problems that those methods cannot. For example, for a system of two variables X and Y, principles like CMC and faithfulness that exploit only conditional and unconditional (in)dependence relations are unable to distinguish between $X \to Y$ and $X \leftarrow Y$. PIM helps because, given various additional assumptions, if it is satisfied by $X \to Y$ then it will not be satisfied by $X \leftarrow Y$, so it can be inferred that the direction in which PIM is satisfied is the correct causal direction (Shimizu et al., 2006).[14] Such

[12]Peters et al. (2017) make this notion of independence precise using algorithmic information theory: they require that $\Pr(X)$ and $\Pr(Y \mid X)$ have zero mutual algorithmic information.

[13]As Woodward (2022a) argues, one of the crucial features of the distinction between laws and initial conditions is that we expect the former to remain stable or invariant under changes in the latter. This invariance requirement is closely connected to the expectation that the mechanisms that generate initial conditions are independent of the laws that evolve those initial conditions.

[14]These additional assumptions take a variety of forms but concern the functional relation between X and Y; for example, that it can be represented by an "additive error model" of the form $Y = f(X) + U$ or $X = f(Y) + V$, or that the distribution of at least one of the variables in the model is non-Gaussian.

methods have been tested on data in which the causal direction is independently known (e.g., the relationship between altitude and temperature) and perform well above chance (Mooij et al., 2016). PIM thus identifies further worldly infrastructure that, when present, can be exploited to facilitate reliable causal inference.

3.6 Realization Independence

It is often the case that upper-level systems of causal relationships display a substantial degree of "realization independence" with respect to their lower-level realizing details: the same upper-level causal relationships continue to correctly characterize the behavior of a system across some range of changes in its lower-level realizers. For example, the thermodynamic behavior of a sample of bulk matter is essentially independent of its exact microstate, as long as that state resides within the appropriate region of the state space. For many psychological and neurobiological phenomena, it appears that the specific behavior of any individual neuron (which is typically stochastic) hardly matters – the aggregate properties of populations of neurons is what matters. Many powerful theoretical tools – such as renormalization group methods and homogenization techniques – were developed to exploit the presence of realization independence, in its myriad forms, to simplify (or make possible at all) prediction and explanation (e.g., (Batterman, 2001)). Realization independence, when it obtains, also allows us to ignore or abstract away from lower-level details (the modeling of which is often intractable) and thus facilitates causal analysis (Woodward, 2018; Robertson, 2021).

If the behavior of some upper-level system exhibits widespread failure of realization independence, then to achieve causal understanding of the system's behavior we must advert to some lower-level description that provides stable, realization-independent relations. If this lower level is not epistemically accessible, or we do not have a good theory of its behavior, or the task of meaningfully connecting it to upper-level behavior is computationally intractable, then our pursuit of causal understanding of the higher-level system will stall. To illustrate the issue, Goldenfeld and Kadanoff (1999) consider the possibility that to adequately model the behavior of a bulldozer, one would have to appeal to quantum chromodynamics. They remark that in this case one would have "model chaos": the choice of a model for bulldozer behavior would be highly sensitive to assumptions about the correct model for the behavior of the strong force, the state of the quark and gluon fields that (partially) realize the bulldozer, etc. This is information that we have no serious possibility of connecting to bulldozer behavior; if it were necessary, bulldozer science would be impossible. If realization independence failed to anything like this degree for sufficiently many natural phenomena, scientific inquiry itself – let alone causal understanding – would be impossible.

3.7 Independence as a Common Thread

All of these infrastructure features concern independence relations: statistical independence among variables, the existence of variables that are causally independent of other variables (making possible interventions and natural exogenous sources of variation), the independence of upper-level relationships from details of their lower-level realizers, the independence of causal relationships from the mechanism(s) determining the cause variable(s) in those relationships (PIM), and the independence of some causal relationships governing the behavior of a system from others (modularity).

This varied store of independence relations provides the primary resources we use to learn about causal relationships: it is the existence of such independence relationships that supports causal learning and the applicability of causal notions. We learn about (causal) dependence, first and foremost, by exploiting information about independence.

4 Does the Infrastructure Call for Further Explanation?

Many of the features that we catalogued as part of the causal infrastructure cry out for further explanation: why do those features obtain in our world? What, for example, accounts for the ubiquity of statistical independencies? Why is it often possible to perform interventions whose effects are largely independent of the lower-level realizers of the manipulated variable? We make no attempt to systematically answer such questions here, but we sketch some scientific details we take to be relevant. We hope this will illustrate how the infrastructure project opens the door to novel lines of inquiry that are easily overlooked when pursuing more familiar metaphysical projects.

Begin with statistical independence. One important consideration concerns certain properties of physical forces in the actual world. Forces between physical bodies decay with distance fairly rapidly: polynomially for the gravitational and electromagnetic forces and exponentially for the weak and strong forces.[15] This means that

[15] At least, at scales larger than about 10^{-5} meters. Inside protons and neutrons the strong force between quarks and gluons increases with distance, resulting in the confinement of those quarks and gluons.

(i) at length scales longer than an atomic radius, only the gravitational and electromagnetic forces are relevant and (ii) at any scale, the dominant contributions to the dynamical evolution of a physical system typically come from other systems in its immediate environment: the decay of forces ensures that interventions on physical systems even relatively short distances away often have negligible effect. (Of course, this depends on the intervention: lighting a candle on the dinner table by striking a match won't affect the temperature of my plate, but lighting it with a flamethrower will.) It is also the case that (iii) essentially all macroscopic systems are electrically neutral under ordinary conditions, which means that electromagnetic interactions will not naturally produce correlations between properties of bulk matter and (iv) the gravitational force is sufficiently weak, and the gravitational effect of any object in our immediate vicinity so dramatically swamped by the gravitational force exerted by the earth, that in many circumstances gravity is also an ineffective means of naturally producing correlations between properties of bulk matter. The result is that many variables of many natural systems, particularly macroscopic systems, will be effectively independent under most ordinary physical conditions.

A second important consideration is that the interaction of a system with a larger environment (or even with other parts of the system itself) can produce decorrelation and effective independence among variables within the system, both classically and quantum mechanically. The environment can, and typically does, interact with the system in such a way as to "wash out" (often rapidly) statistical correlations within the system. For example, although two molecules in a gas will have correlated momenta immediately after they collide, their momenta rapidly become effectively independent as a result of subsequent independent collisions of each particle with other molecules in the gas. In quantum theories, environmental decoherence produces a similar result. Consider a "system" – two particles, with entangled spins – interacting with molecules constituting an ambient "environment". This interaction rapidly entangles the system with the environmental molecules, and although the entanglement between the two initial spins entails that results of appropriate measurements of those spins should be correlated, the entanglement between system and environment renders these correlations inaccessible to us via measurements. The result is that environmental decoherence rapidly renders measurements of the spins of the "system" particles effectively statistically independent (see e.g., (Schlosshauer, 2019)).

Another important consideration, relevant to the above, is that whether variables are independent can depend on their "grain" as well as the "cut" we make to distinguish the system of interest from some larger environment. By choosing an appropriate graining of variables and an appropriate system/environment cut, we can sometimes characterize systems in such a way that few variables are correlated and many are independent, thus maximizing the information that can be exploited for causal discovery. In particular, a description of a system using fine-grained variables which are dependent can sometimes be replaced by a description using more coarse-grained variables that are independent. For example, the exact positions of atoms in my coffee cup may be, at some divinely accessible level of precision, correlated with my position at my desk because of gravitational interactions. But if we adopt more tractable, coarse-grained variables – variables that track atomic positions to only the 10th decimal place would suffice, let alone variables that track only the position of the coffee cup itself on the desk – there will be no such correlation. (Incidentally, this is one reason why the claim that everything in the backward light cone of some event E is causally related to E shows much less than is often supposed. The claim might be correct, in some sense, for maximally fine-grained descriptions of all systems and events involved, but even mildly coarse-grained variables will exhibit effective independence. It is almost always those variables that are of interest, even in the practice of fundamental physics (cf. Frisch, 2014, pp. 68-70).[16]

Causal inference does not require that variables (or relationships) be fully independent (or modular) at all spatiotemporal scales, but only that they be effectively independent at a particular scale of interest. To claim that rainfall is exogenous with respect to crop growth (Simon and Rescher, 1966), it need not be that agriculture has no long-term influence on climate. It only need be that over the timescale of interest (say, five years) any influence of crop growth on rainfall is sufficiently small to be negligible on that scale. That independence is often only approximate is not in tension with recognizing independence as a genuine feature of the world. It is still the case that the possibility of differentiating between dependent and independent variable sets at one scale requires it not to be the case that all variables influence all others equally at all scales (Weinberger, 2020). It is fortunate for causal analysis that it is possible to successfully characterize the dominant behaviors of worldly systems at different scales in terms of a relatively small set of causal influences.

5 Success, Realism, and Effective Theories

We have argued that the presence of the worldly infrastructure features explains the success of our procedures for making reliable inferences about causal relation-

[16] It is worth adding that when variables are dependent, transformations of those variables may be available that yield independent variables. All else being equal, we tend to prefer the latter. If q is position and p is momentum, then the variables $U = q + p$ and $V = q - p$ are statistically dependent, but q and p will generally not be. This is one reason for preferring q and p over U and V.

ships.[17] We view this argument as an instance of a general form of reasoning employed throughout science. Scientific modeling rests on empirical presuppositions. The application of textbook techniques for calculating scattering cross sections in quantum field theory involves a number of them: for example, that the scattered particles interact weakly and that the scattering takes place at effectively zero temperature. The use of polygenic risk scores to predict behavioral traits and diseases rests on the presupposition that the traits in question are the result of a large number of genetic factors, each individually contributing a very small and approximately additive effect on the trait in question; it also presumes that the ingredients (SNPs) that go into the scores, even if not themselves causal, are correlated with genuinely causal factors (Kendler and Woodward, n d). When these applications are successful, we can infer that the real-world physical situation satisfies the presuppositions of our modeling strategies. For example, the fact that polygene risk scores are replicable and predictively successful suggests to geneticists that these presuppositions concerning genetic architecture are correct. It also helps to explain why many previous attempts to identify common candidate genes with large effects on the traits in question were unsuccessful (in the sense of failing to replicate): for many traits, such genes do not exist. Concluding that the presuppositions of a modeling strategy are satisfied in the real world on the basis of that strategy's empirical success is a familiar inference in the sciences. We claim that a similar inference holds for the worldly infrastructure that underlies successful causal reasoning: the success of various causal discovery strategies reveals facts about what the world is like – i.e., about the presence of the worldly infrastructure that supports the success of those strategies. In this sense, we are simply treating causal discovery like other fields of successful scientific inquiry.

[17]More precisely, it is an indispensable part of the explanation. The success of those procedures also depends on obvious factors like the existence of users of the procedures, their cognitive abilities, etc.

This "argument from success" may remind readers of similar arguments by scientific realists: (1) science (or a particular theory) is predictively accurate, (2) the success of science (or this particular theory) would be miraculous if its descriptions of the world were not at least approximately true, therefore (3) those descriptions are at least approximately true. We think our argument from success is stronger than this familiar argument in several respects.

One common objection to this argument for scientific realism invokes the specter of underdetermination. One way to state the objection is that no matter the predictive success of $\mathbf{T}$, one cannot rule out that $\mathbf{T}$ is false and some alternative $\mathbf{T}^*$ is true, where $\mathbf{T}^*$ is inconsistent with $\mathbf{T}$ but accounts for the same phenomena. How seriously one should take this objection in a given context depends, in part, on how plausible it is that there is any such theory $\mathbf{T}^*$. Applied to our argument, the analogous issue is whether there is some alternative account, postulating very different worldly infrastructure, that accounts equally well for the "phenomena": the success of our causal inference procedures. We are not aware of even a hint of a serious proposal in this direction.

A second disanalogy: our reasons for thinking that the worldly infrastructure features are present go well beyond the inference to the best explanation on which the scientific realist relies. Often we can directly establish that the infrastructure features *are* present – we observe that certain variables are statistically independent, etc. Direct observation that the infrastructure features are present, combined with arguments that *if* present, they would explain the success of our causal discovery procedures, *and* that no other potential explanations of success seem to be available, all support the claim that the presence of the infrastructure features is crucial for the success of causal discovery. The resulting argument is thus stronger than the familiar inference to the best explanation of the scientific realist. We have something more like an argument that the presence of the infrastructure features would be the *only* empirically viable explanation, bolstered by independent evidence that the potentially explaining features are, in fact, present.

We conclude by commenting on another issue that, although distinct, might invite a similar skeptical challenge: the status of the causal claims that our inference procedures deliver. Suppose those procedures seem to tell us that (1) C causes E. How do we know that some alternative account – perhaps a more "fundamental", fine-grained account – won't one day be accepted, according to which (1) is not true, but rather some alternative (2) is true: C^* causes E? If so, our causal inference procedures will not be successful in the sense we have claimed.

One way this might happen is if there was some unrecognized source of confounding in the procedures that we took to establish (1): we thought we were intervening on C and observing changes in E, but some other factor Z was associated with our attempted interventions and in fact Z causes E. This is always possible in principle, but is extremely far-fetched in many realistic cases and often not a possibility to be taken seriously in inquiry.

Suppose instead that we succeed in performing genuine interventions on C and that E changes in accord with (**M**), or we provide evidence via other inference strategies we have discussed for the same conclusion about how E would respond to interventions on C. Might we still subsequently discover that (1) is false? This is harder to envision. This could be a genuine possibility if (**M**) was an inadequate account of causation, with the correct account either one that did not incorporate (**M**) but instead invoked some entirely distinct condition(s) S, or one according to which a relationship is causal only if it satisfies (**M**) *and* some additional condition R that (**M**) omits. Then it might be that even though E changes in response to interventions on C, the condition(s) S or R are not satisfied so the relation between C and E is not really causal.

Of course, if (**M**) or something sufficiently similar is the correct account of causation, that won't be a genuine possibility. However, there is reason beyond optimism about (**M**) to believe that statements of the form "interventions on C are associated with changes in E" are unlikely to be overturned by future developments. Statements of that form, established on the basis of apparently well-designed experiments and reliable inference procedures but then invalided by subsequent developments, are difficult to come by. Subsequent discoveries may tell us more about why interventions on C are associated with changes in E, about which components of C are the genuine difference-making elements, about mediating variables, about the range of conditions under the original claim holds, and so on. However, they rarely show that under the specified conditions the original claim is false. (And if they do, the mistake is usually discoverable via investigation at the level of the original causal claim, such as the presence of an unknown confounder or a failure of replicability, and not via some means that depend on future theorizing about causation.)[18] This is, in part, because interventionist causal claims are "thin", i.e., relatively ontologically non-committal. In particular, interventionist causal claims describe a dependence relationship without making commitments about more "fundamental" or fine-grained characterizations of the variables involved, nor about the specific mechanism responsible for dependence

[18] For example, one way that a causal claim might be mistaken is that the cause variable C might be discovered to be "ambiguous": different interventions that set C to the same value might be associated with different outcomes depending on how C is "realized" at some lower level (Spirtes and Scheines, 2004). This mistake is something that can typically be discovered by doing experiments at the level of C – no discovery of a deeper or more fine-grained theory is required.

relationship. For example, if we understand "aspirin causes headache relief" as a "thin" interventionist claim, it is wildly implausible that it would be falsified by subsequent theoretical developments. After all, it makes no commitments about how the causal influence is transmitted (e.g., via inhibiting prostaglandin production vs. via activating μ-opioid receptors) nor about the chemical composition of aspirin, the physiological realizers of headaches, etc. This results in interventionist relations typically being preserved under theory change, including the embedding of high-level causal relations into more "fundamental" or fine-grained theoretical descriptions.

This illustrates another important disanalogy with the familiar argument for scientific realism. There are many examples of theories that are highly predictively successful but have mistaken ontological commitments –nineteenth-century theories that held light and electromagnetism were transmitted by a mechanical ether, Dirac's "hole theory" that was used to predict the existence of antimatter, and so on. By contrast, statements of the form "interventions on C are associated with changes in E" describe a dependence relationship without making such ontological commitments. This makes it possible for interventionist causal claims to survive subsequent scientific developments and attendant changes in ontology. This is just as true for claims that more than 40 CAG repeats in the HTT gene causes Huntington's chorea, that the motion of a conducting material in a magnetic field causes induction of a current, etc. as it was for the claim that aspirin causes headache relief.

Our invocation of "well-designed" experiments and inference procedures may seem question-begging. We disagree. One can always be mistaken in thinking that an experiment is well-designed, but whether this is the case is usually something that can be determined by subsequent scrutiny of the experiment itself. Consider the replication crisis affecting portions of psychology. That the experiments claiming to show that certain manipulations cause various effects (e.g., priming effects on behavior) were badly designed can be revealed from inspection of the experiments themselves and the techniques employed to analyze them (e.g., inappropriate statistical procedures), along with the information provided by their failure to replicate. Again, we can discover this via investigations at the same level as the original causal claim; new scientific theories are not required. Our point is not that this is never the case – nineteenth-century experimenters working with cathode ray equipment could not possibly have recognized that their experiments were poorly designed and that certain causal claims rested on shaky ground (because they failed to control for X-rays) until X-rays were discovered by Roentgen in 1895. Our point is that such cases are the exception, not the rule, and do not justify anything like a general expectation that interventionist claims are endangered by future scientific development.

Physicists make increasing use of the notion of an "effective theory". This is a theory T_E that accurately captures dependence relations over a restricted range of scales or within a specified domain, but will not be predictively accurate outside that domain. The use of effective theories is licensed by the fact that their structure is relatively independent of the structure that might be revealed by more fine-grained levels of analysis. Put otherwise, the dependence relationships described by T_E depend only weakly on which fine-grained theory turns out to be correct: as long as that fine-grained theory lies in a class of theories that satisfy specified conditions, the relationships in T_E will be retained by that theory (Williams, 2019).[19] (In physics, this can often be demonstrated by the methods for exploiting realization independence mentioned in section 3.6 or similar techniques.) The use of effective theories in particle physics has received philosophical attention, but the concept can be profitably applied

[19]The formal specification of these conditions varies between theories, but they are in general fairly weak.

more generally; for example, the use of the Navier-Stokes or Navier-Cauchy equations to model the continuum-scale properties of fluids and solids (respectively) can be illuminatingly understood as making use of effective theories (Batterman, 2021). We suggest that causal claims, in an interventionist framework like (**M**), exhibit many of the same characteristics as the dependence relations found within effective theories throughout physics: they are typically independent of many details of their lower-level realizers while remaining non-committal about any detailed ontological account of those realizers. As discussed above, these properties justify the expectation that interventionist claims will remain stable across future theoretical developments.

6 Disanalogies with Traditional Metaphysics

We presume it is clear that the "worldly infrastructure" project differs from those typically engaged in under the heading of the metaphysics of causation. As promised in the introduction, here we spell out these differences in more detail, with the aim of further elucidating our approach.

No Special Entities or Relationships: The characterization of the infrastructure does not require special metaphysical concepts or entities: no powers, no relations between universals, no Humean mosaic of perfectly natural properties, etc. Instead, a characterization of the worldly infrastructure supporting causal reasoning can be given using non-metaphysically-loaded notions that are already employed by the workaday scientist or mathematician (including disciplines like statistics and econometrics) – notions like statistical independence and exogeneity.

Links to Epistemology: We take it to be a guiding principle of the natural sciences that successful methods for detecting empirically accessible manifestations systems in the natural world are not independent of what those systems are actually like. Methods for detecting genes and investigating their behavior are not independent of features that genes possess and general facts about genetic architecture. Similarly for methods for detecting elementary particles. We think that causation should be approached in a broadly similar way. By carefully attending to the epistemological successes of frameworks for causal modeling and inference, one can extract a number of conclusions about the worldly infrastructure that supports causal reasoning. Unlike more familiar metaphysical discussions of causation, the infrastructure project maintains tight connections with the epistemology and methodology of causal inference since the infrastructure features are exactly those on which successful causal inference relies.

Metaphysical Necessity: Metaphysicians of causation often understand their task as characterizing features of causation that are "metaphysically necessary" – perhaps features that causation must possess in all "metaphysically possible" worlds. The infrastructure project has more restrained modal ambitions. It does not claim that the generic features that characterize the worldly infrastructure obtain as a matter of metaphysical necessity, nor does it claim that there is a metaphysically necessary connection between causation and the presence of those features. It does not attempt to identify causation with the infrastructure features and it does not take the infrastructure features to be "grounds" or "truth makers" for causal claims, in the metaphysician's sense of those terms (which imports assumptions about metaphysically necessary relations). Instead, the infrastructure project is solely concerned with properties of the actual world and how they support causal reasoning. In line with its restricted modal ambitions, the infrastructure project declines to take a stand on how to think about causal relations in "alien worlds" – proposed worlds very different from our own in which the worldly infrastructure is not present or not connected with causation in anything like the way it is in the actual world. While entertaining such possibilities may be worthwhile for those interested in identifying metaphysically necessary features of causation, they lie outside the purview of an attempt to characterize the (actual) worldly infrastructure of causation.

Appeal to intuitions: Relatedly, the infrastructure project is not an attempt to systematize intuitions about causation. In this regard, it diverges both from the standard methodology for studying actual causation (Paul and Hall, 2013) and the methodology commonly employed when using judgments about non-actual worlds, especially "alien worlds", to extract lessons about causation and causal reasoning. The alien worlds typically considered are worlds in which, by construction, some or all of the worldly infrastructure that supports successful causal reasoning in the actual world is missing. It is no surprise, then, that we do not consider the judgments one may be tempted to make about such worlds – using causal concepts and reasoning strategies developed in the actual world to exploit precisely that worldly infrastructure – to provide useful information or insight into causation or causal reasoning in the actual world. Indeed, our functionalism inclines us toward skepticism about whether there is a determinate fact about how causal concepts developed to exploit a specific worldly infrastructure can be applied to worlds that do not exhibit the infrastructure presupposed by those concepts and strategies.[20]

Domain-specificity: Although the infrastructure features obtain generically, it is entirely possible that there are actual systems that, at least at some natural levels of analysis, fail to instantiate the infrastructure features to such an extent that they

[20]For additional discussion, see (Williams, 2022b; Woodward, 2022b).

resist causal analysis. For example, there has been recent discussion about whether structural features of general relativity leave it unable to provide causal explanations of various phenomena. Several of those structural features apparently imply that certain interventions (and the associated counterfactuals) that are crucial for causal interpretation are ill-defined, which would mean a crucial element of the worldly infrastructure is missing.[21]

If this is correct, and if sufficiently many other infrastructure features are absent for certain systems modeled in GR, we think it warranted to conclude that the behavior of such systems simply will not admit a straightforward causal interpretation, at least on anything like how we presently think about causation. This possibility illustrates our general point that some systems – at least when modeled at certain levels of analysis – may simply be "unfriendly" to causal analysis because important worldly infrastructure is not present. Accepting this contingency of the worldly infrastructure stands in contrast to the common philosophical expectation that causal notions should be applicable everywhere that certain minimal constraints (e.g., the presence of regularities) are satisfied.

Relevance of Physics: The causal modeling techniques that inform our analysis are widely and successfully used in the social and behavioral sciences and increasingly in portions of biology, including neurobiology and genetics. However, readers may

[21]For discussion, see e.g., (Curiel, 2015; Jaramillo and Lam, 2021). Problems with the relevant counterfactuals arise in several ways. The stress-energy tensor is partly dependent on metric structure for its characterization, so an intervention on the former with respect to the latter is arguably not well-defined. This creates difficulties for the claim that any chosen matter distribution "causes" metric structure. In addition, the absence of unique vacuum solutions creates problems for counterfactuals concerning what would happen if matter were removed from a region of spacetime (since there won't be unique situations associated with the antecedents of such counterfactuals). There are also global constraints that follow from the field equations that create problems for the possibility of local interventions. We emphasize that we don't think these are problems for our project; rather these pinpoint features of GR that make causal interpretation challenging.

wonder whether these techniques can be usefully applied within physics and, relatedly, whether the infrastructure features we describe are important for causal reasoning in physics. We have several replies. First, there is now a rapidly growing literature applying causal discovery methods to classical and quantum physics, e.g., (Costa and Shrapnel, 2016; Janzing et al., 2016; Allen et al., 2017; Barrett et al., 2020). Perhaps it is unsurprising, then, that the infrastructure features on which we have focused do often figure in physical reasoning. For instance, assumptions that causally unrelated variables will be statistically independent very common; for example, this is why, in the absence of delicately arranged contrivances, the possibility of coherent electromagnetic radiation converging on a source it is not considered a serious physical possibility in the actual world, despite being consistent with Maxwell's equations. The need for theoretical structure to license interventions also plays an important (albeit rarely explicit) role in physical theorizing, from thermodynamics to quantum field theory.[22] The principle of independent mechanisms is also universally assumed to hold in physics and has been put to use in classical and quantum physics for multiple purposes (e.g., Maudlin, 2007, pp. 130-35; Janzing et al., 2016; Williams, 2022a).

Second, to the extent that these infrastructure features are not satisfied by some physical systems and we cannot usefully apply causal inference techniques, we think the appropriate conclusion is that this may simply mark limits on the applicability of causal thinking in physics. However, we do not conclude from these limits (should they exist) that the features we discuss are somehow thereby unimportant to causal

[22]Treating thermodynamics as a "control theory" is a familiar idea (Wallace, 2014), especially in engineering, but the invocation of quantum field theory may be surprising. The cluster decomposition principle, mentioned above, is one of multiple principles that plays this role: it ensures that particles separated by sufficiently large spatial or temporal intervals can be treated as dynamically independent. This means that one can intervene on the states of particles for example, when representing the preparation of particles prior to a scattering experiment in a way that is crucial for representing scattering theory.

reasoning in circumstances in which they are present.

Primacy of Physics: It is fairly often supposed (and occasionally argued) that the legitimacy of causation within science depends on whether causal relations figure in the ontology of fundamental physics. The infrastructure project does not accord physics privileged status of this sort. The legitimacy of causal reasoning in genetics does not hinge on whether the Standard Model of particle physics provides causal explanations of phenomena; whether and when causal reasoning can be applied to any domain of natural phenomena, including physics itself, depends on whether the worldly infrastructure that supports causal reasoning is present.

Sufficiency of metaphysical concepts: We have emphasized that many concepts that are commonplace in metaphysical analyses of causation are unnecessary for characterizing the worldly infrastructure that causal inference exploits. In fact, the concepts in that familiar metaphysical toolkit also may not *suffice* for characterizing a causal system. For example, a number of philosophers hold that there is a tight connection of some type between causal relations and laws of nature. A simple and very strong form of this connection might assert that C causes E if and only if C and E "instantiate", or are "related by", a "law of nature". We do not necessarily object to the "only if" direction of this claim, depending on how the terms in quotation marks are made precise. It is, however, our view that instantiating the worldly infrastructure required for causal reasoning requires considerably *more* than the satisfaction of the "if" direction. The infrastructure features include facts about the presence of statistical independencies, the possibility of interventions, and so on. If these features do not obtain then it may not be true that C causes E, even if C and E are "related by" a "law of nature". Our discussion of obstacles to defining interventions and their associated counterfactuals in general relativity provide an example; anti-entropic systems also appear to defy causal description, despite obeying the same laws of nature as entropic systems (Williams, 2022b). These examples illustrate the general point that more is required for causal notions to be applicable to a given system or relationship than merely that it be lawlike.

7 Conclusion

This paper engages in a novel project – the elucidation of the worldly infrastructure that supports the application of causal analysis. This is distinct from most projects pursued within mainstream metaphysics of causation, but it also does not concern only the epistemology or methodology of causal reasoning. Instead, it aims to identify the features that are "out there" in nature that underlie our ability to learn about causal relationships and successfully apply them. The project is is motivated by an idea that is relatively uncontroversial elsewhere in the philosophy of science: when a theory or methodology delivers reliable knowledge about its subject, it is worthwhile to investigate the worldly features that help to explain that success.

References

Allen, J. M. A., Barrett, J., Horsman, D. C., Lee, C. M., and Spekkens, R. W. (2017). Quantum common causes and quantum causal models. *Physical Review X*, 7(3):031021.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Barrett, J., Lorenz, R., and Oreshkov, O. (2020). Quantum causal models. *arXiv preprint: 1906.10726*.

Batterman, R. (2001). *The devil in the details: asymptotic reasoning in explanation, reduction, and emergence*. Oxford University Press.

Batterman, R. (2021). *A Middle Way: A non-fundamental approach to many-body physics.* Oxford University Press.

Brown, H. R., Myrvold, W., and Uffink, J. (2009). Boltzmann's h-theorem, its discontents, and the birth of statistical mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 40(2):174–191.

Cartwright, N. (1999). Causal diversity and the markov condition. *Synthese*, 121(1/2):3–27.

Costa, F. and Shrapnel, S. (2016). Quantum causal modeling. *New Journal of Physics*, 18(6):063032.

Curiel, E. (2015). If metrical structure were not dynamical, counterfactuals in general relativity would be easy. *arXiv preprint: 1509.03866*.

Dougherty, J., Miller, M., and Williams, P. (2023). The aim and structure of cluster decomposition. *manuscript*.

Elliott, K. and Lange, M. (2022). Running it up the flagpole to see if anyone salutes: a response to woodward on causal and explanatory symmetries. *THEORIA: An International Journal for Theory, History and Foundations of Science*, 37(1):53–62.

Frisch, M. (2014). *Causal reasoning in physics*. Cambridge University Press.

Glymour, C. (2006). Markov properties and quantum experiments. In Demopoulos, W. and Pitowsky, I., editors, *Physical Theory and Its Interpretation*, pages 117–125. Springer Dordrecht.

Goldenfeld, N. and Kadanoff, L. (1999). Simple lessons from complexity. *Science*, 284(5411):87–89.

Greenspan, R. (2001). The flexible genome. *Nature Reviews Genetics*, 2:383–387.

Hausman, D. M. and Woodward, J. (1999). Independence, invariance and the causal markov condition. *The British Journal for the Philosophy of Science*, 50(4):521–583.

Hernán, M. A. and Robins, J. M. (2020). *Causal Inference: What If*. Chapman & Hall/CRC.

Janzing, D., Chaves, R., and Schölkopf, B. (2016). Algorithmic independence of initial condition and dynamical law in thermodynamics and causal inference. *New Journal of Physics*, 18(9):093052.

Jaramillo, J. L. and Lam, V. (2021). Counterfactuals in the initial value formulation of general relativity. *The British Journal for the Philosophy of Science*, 72(4):1111–1128.

Kendler, K. and Woodward, J. (n. d.). Polygenic risk scores. *manuscript*.

Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N., and Smith, G. D. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in medicine*, 27(8):1133–1163.

Maudlin, T. (2007). *The metaphysics within physics*. Oxford University Press.

Mitchell, S. (2009). *Unsimple Truths*. University of Chicago Press.

Mooij, J. M., Peters, J., Janzing, D., Zscheischler, J., and Schölkopf, B. (2016). Distinguishing cause from effect using observational data: methods and benchmarks. *Journal of Machine Learning Research*, 17(1):1103–1204.

Paul, L. A. and Hall, N. (2013). *Causation: a user's guide*. Oxford University Press.

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

Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference: foundations and learning algorithms*. The MIT Press.

Robertson, K. (2021). Autonomy generalised; or, why doesn't physics matter more? *Ergo*, forthcoming.

Rubin, D. B. (1974). Estimating causal effects in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688.

Sanderson, E., Glymour, M. M., Holmes, M. V., Kang, H., Morrison, J., Munafò, M. R., Palmer, T., Schooling, C. M., Wallace, C., Zhao, Q., and Davey Smith, G. (2022). Mendelian randomization. *Nature Reviews Methods Primers*, 2(1):1–21.

Schlosshauer, M. (2019). Quantum decoherence. *Physics Reports*, 831:1–57.

Shimizu, S., Hoyer, P. O., Hyvärinen, A., Kerminen, A., and Jordan, M. (2006). A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7(10):2003–2030.

Simon, H. A. (1953). Causal ordering and identifiability. In Hood, W. and Koopmans, T., editors, *Studies in Econometric Method*, pages 49–74. Wiley.

Simon, H. A. and Rescher, N. (1966). Cause and counterfactual. *Philosophy of Science*, 33(4):323–340.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, prediction, and search.* The MIT Press.

Spirtes, P. and Scheines, R. (2004). Causal inference of ambiguous manipulations. *Philosophy of Science*, 71(5):833–845.

Wallace, D. (2014). Thermodynamics as control theory. *Entropy*, 16(2):699–725.

Weinberg, S. (1995). *The Quantum Theory of Fields, Volume I: Foundations.* Cambridge University Press.

Weinberger, N. (2020). Near-decomposability and the timescale relativity of causal representation. *Philosophy of Science*, 87(841-856).

Williams, P. (2019). Scientific realism made effective. *The British Journal for the Philosophy of Science*, 70(1):209–237.

Williams, P. (2022a). Entanglement, complexity, and causal asymmetry in quantum theories. *Foundations of Physics*, 52(2):1–38.

Williams, P. (2022b). The fate of causal structure under time reversal. *THEORIA: An International Journal for Theory, History and Foundations of Science*, 37(1):87–102.

Wood, C. J. and Spekkens, R. W. (2015). The lesson of causal discovery algorithms for quantum correlations: causal explanations of bell-inequality violations require fine-tuning. *New Journal of Physics*, 17(3):033002.

Woodward, J. (2005). *Making things happen: a theory of causal explanation*. Oxford University Press.

Woodward, J. (2014). A functional account of causation; or, a defense of the legitimacy of causal thinking by reference to the only standard that matters – usefulness (as opposed to metaphysics or agreement with intuitive judgment). *Philosophy of Science*, 81(5):691–713.

Woodward, J. (2018). Explanatory autonomy: the role of proportionality, stability, and conditional irrelevance. *Synthese*, 198(1):237–265.

Woodward, J. (2022a). Flagpoles anyone? causal and explanatory asymmetries. *THEORIA: An International Journal for Theory, History and Foundations of Science*, 37(1):7–52.

Woodward, J. (2022b). Responses. *THEORIA: An International Journal for Theory, History and Foundations of Science*, 37(1):111–129.