The Mathematics of Causal Capacities

David Danks

Departments of Philosophy & Psychology

Carnegie Mellon University

ddanks@cmu.edu

Abstract

Models based on causal capacities, or independent causal influences/mechanisms, are widespread in the sciences. This paper develops a natural mathematical framework for representing such capacities by extending and generalizing previous results in cognitive psychology and machine learning, based on observations and arguments from prior philosophical debates. In addition to its substantial generality, the resulting framework provides a theoretical unification of the widely-used noisy-OR/AND and linear models, thereby showing how they are complementary rather than competing. This unification helps to explain many of the shared cognitive and mathematical properties of those models.

1. Introduction

In many scientific domains, one finds models focused on causal influences that function (at least somewhat) independently of one another. For example, cognitive models are typically expressed in terms of distinct cognitive processes that have no direct influence on one another's functioning, and so can proceed independently, whether sequentially or in parallel. As just one instance, many theories of categorization posit that people first perceive the relevant stimulus, then judge its similarity to various known categories, and finally use those similarity judgments to generate a behavioral response. These processes obviously matter for one another; the output of the perceptual process, for example, is the input to the similarity judgment process. But in essentially all similarity-based cognitive theories of categorization, the *functioning* of one process is assumed to be largely independent of the *functioning* of the other processes. The "inner workings" of the perceptual process are assumed to be irrelevant to the way that similarity judgments are made; the only influence of the former on the latter is the particular information that it outputs.

More generally, scientific models and theories frequently divide the world into distinct processes (typically, causal ones) such that the operation of one process has minimal dependence on—in the best case, true independence from—the operations or states of other processes. Probably the clearest articulation of this picture is based on the notion of causal capacities (Cartwright 1989, 1999, 2007; Martin 2008; see also Heil 2005), but similar ideas can be found in many writings on mechanisms (in the spirit of Machamer, Darden, & Craver, 2000). In this paper, I focus on such independent causal influences; for convenience, I will refer to them as 'capacities', but this term should be understood broadly. The basic idea is that capacities are just those causal powers that a cause *C* has purely by virtue of being a *C*; causal capacities are

"something they [the causes] can be expected to carry with them from situation to situation" (Cartwright 1989, 145). That is, capacities inhere in *C* rather than arising from the particular situation, and so their operation should be relatively unaffected by other processes in the system. This (almost) independence is exactly what enables the construction of "nomological machines" (Cartwright 1999, 2007) that generate the regularities—some contingent, some law-like—that we observe and manipulate.

The philosophical literature on causal capacities and mechanisms has largely focused on questions that are metaphysical (e.g., are they basic/fundamental features of the world?) or epistemological (e.g., can we discover capacities from observational or experimental data?). I here consider a representational question: is there a natural, privileged representational framework for systems in which the causal influences[1] are independent[2] of one another (i.e., each does not depend on the values, operations, or status of the others)? There is enormous variety in the world, and so any representational framework inevitably simplifies or is sometimes not applicable. My interest here is in a representational framework that applies to the "standard" or "ordinary" cases, and so can function as a default framework; I use the terms 'natural' and 'privileged' to refer to such a framework. One might think that there obviously can be *no* such

[1] For simplicity, I assume that each independent influence corresponds to a single cause, as multiple (interactive) causes can be merged into a single, multidimensional, factor.

[2] This independence should not be confused with (a) statistical independencies that can be used to (sometimes) infer causal structures from data (Spirtes, Glymour, and Scheines 2000); or (b) the idea of 'modularity' to refer to causal connections that can be separately intervened upon (Hausman and Woodward 1999, 2004; Cartwright 2002).

privileged representation, as the independence property seems too weak for this task, but that response turns out to be mistaken.

2. The Mathematics of (a Special Case of) Causal Capacities

2.1. The Noisy-OR/AND Model

Assume that we have a set of (possible) causes $C_1, \ldots, C_n$ and a target effect E. The functioning of C_i's capacity is supposed to inhere in C_i, and so the causal strength or influence of C_i should be representable without reference to the states of the other variables. In particular, C_i's impact on E should not depend on the state or causal strength of C_j, and it should be monotonic in C_i; in particular, even if the quantitative impact is not constant across values of E (due to, e.g., saturation of E), the valence should not depend on E's value. Finally, for mathematical tractability, I assume that each variable's possible values can be represented as numbers, though each variable can have its own scale; this is a trivial assumption when the variables are binary (i.e., two-valued), but is non-trivial in other cases (e.g., there is no privileged way to map *red*, *green*, and *blue* to numbers).

Consider the special case situation in which all factors—causes and the effect—can be represented as binary variables. For this case, a privileged mathematical framework (with origins in 19th century mathematics) has been developed in machine learning and cognitive psychology (Good 1961; Srinivas 1993; Heckerman and Breese 1994, 1996; Cheng 1997; Glymour 1998; Cozman 2004). Suppose that we have a single generative (binary) cause C_1 of the (binary) effect E, and so E occurs when (and only when) C_1 is present and the capacity of C_1 is active, where w_1 is the strength of that capacity. Thus, we immediately derive $P(E) = w_1 \times \delta(C_1)$, where $\delta(X) = 1$ if X is present, 0 if X is absent. If we have a second generative cause C_2 of E, then E occurs when

(and only when) either C_1 or C_2 generates it, where the 'or' is non-exclusive. Thus, we have $P(E) = w_1\delta(C_1) + w_1\delta(C_2) - w_1\delta(C_1)w_2\delta(C_2)$; that is, the probability of E is just the sum of the probabilities that it is caused by one cause, minus the probability that both caused it (in order to account for that case being "double-counted" in the sum of the first two terms). More generally, if we have n distinct, independent generative causes, then the resulting expression for $P(E)$ is the "noisy-OR" model (Good 1961; Kim and Pearl 1983; Pearl 1988; Srinivas 1993; Heckerman and Breese 1994; Cheng 1997; Glymour 1998):

$$P(E|C_1, \ldots, C_n) = 1 - \prod_{i=1}^{n}(1 - w_i\delta(C_i)) \tag{1}$$

In a noisy-OR model, E is an OR-function of the different causes, but with cause-specific "noise" (understood instrumentally) that probabilistically makes that cause's capacity inactive. Thus, the probability that E occurs is just the probability that at least one present cause has an active capacity. Moreover, equation (1) is uniquely privileged: it is the *only* equation for purely generative binary causes with distinct causal capacities (i.e., independent causal influences) that satisfies various natural properties (Cozman 2004).

Of course, not all causes are generative; we are often interested in causes that *prevent* the effect from occurring. If a preventive cause P interferes with the functioning of only one specific generative cause G,[3] then P has the (mathematical) impact of reducing G's causal strength and so we can combine their causal capacities. We cannot do the same for preventers that apply to all generators equally; such preventers operate as (noisy, probabilistic) "switches" that control whether any generative cause can be active at all. That is, E occurs when (and only when) at least

[3] An ambiguity lurks here between "prevention as blocking" and "prevention as reducing," but I postpone discussion of this ambiguity until later in this section.

one generative cause's capacity is active and none of the preventive causes' capacities is active. This relationship is captured by a "noisy-OR/AND" model, since the generative causes combine in a noisy-OR function, whose result is then combined with a noisy-AND function for the preventive causes (i.e., the effect occurs only if a generator is active AND P_1 is not active AND … P_m is not active):

$$P(E|C_1,\ldots,C_n,P_1,\ldots,P_m) = \prod_{j=1}^{m}(1-w_i\delta(P_j))\Big[1-\prod_{i=1}^{n}(1-w_i\delta(C_i))\Big] \quad (2)$$

This equation provides (arguably) the most natural representation of causal capacities, both generative and preventive, that exert independent causal influence (Srinivas 1993; Heckerman and Breese 1994, 1996; Lucas 2005). Moreover, there is substantial empirical evidence that humans preferentially represent causal systems as functioning according to equation (2) (Cheng 1997; Holyoak and Cheng 2011; Danks 2014).[4]

2.2. Resolving Ambiguities

Although there is great value in this mathematical framework, the restriction to binary variables is significant, as there are many cases in which the influence of a causal capacity depends in part on the factor's magnitude or intensity, or the effect can exhibit fine degrees of meaningful variation. Before generalizing the noisy-OR/AND model to many-valued variables, however, we must clarify two key conceptual (though not mathematical) ambiguities.

[4] The connection between psychological theory and capacities is unsurprising, as Cheng's (1997) causal power theory in cognitive psychology was explicitly modeled on Cartwright's (1989) capacity account of causation.

Mathematically speaking, binary variables are simply those with two possible values. When talking about causal capacities, however, a more specific interpretation is typically intended: factors can be "present" vs. "absent" or "on" vs. "off"; capacities can be "active" vs. "inactive". These interpretations provide a natural value ordering, as shown by the standard practice of mapping "present" to the value of 1 and "absent" to the value of 0.[5] More generally, we typically understand the "absent" or 0 value to be the *lower bound* of the possible values for that variable. At the same time, the zero value in the context of causal capacities almost always serves as the *baseline* value: it is the value that E would have if nothing influenced it. This second role of the zero value is clear in the mathematics of the noisy-OR/AND model, as $P(E = 0 \mid$ all generative causes are absent$) = 1$. That is, the standard model of (binary) causal capacities assumes that absence is the appropriate "uncaused" state for E.[6]

These two different roles for zero—lower bound and baseline value—are conceptually distinct and empirically distinguishable. For example, in most terrestrial environments, the baseline value for *Oxygen in Room* (i.e., the value it has when represented causes are all inactive) is "present," not "absent." We can represent this different baseline value in the noisy-OR/AND model, but only through a mathematical trick (namely, a very strong, always-present generative cause). A better solution would be to allow the lower bound and baseline to diverge. This

[5] This particular mapping could obviously be reversed without any change in substantive content, though 'lower bound' and 'upper bound' would need to be swapped in what follows.

[6] One might worry that "uncaused states" are impossible. However, if causes function independently, then it is at least theoretically possible for none to be active at a moment in time. More generally, *any* model with independent causal influences yields a baseline value, even if it is only ever theoretically (rather than empirically) realized.

generalization does not matter for cases with only binary variables, as any model with variables whose baseline is 1 can be translated into a model in which all baselines are 0. Outside of this special case, however, the baseline value plays a distinct mathematical role, and so any model of causal capacities that allows for more-than-binary variables (such as the one developed in Section 3) must distinguish conceptually between the lower bound value and the baseline value.

The multiple roles played by zero point towards the other important ambiguity in the standard noisy-OR/AND model of causal capacities. Because the zero value is both the lower bound and the baseline, there are two different ways to prevent, or make *E* less likely. First, the preventer could stop generative causes from exerting their usual influence. These *blockers* serve to keep the effect variable closer to its baseline value, as they (potentially) eliminate causal influences that drive the effect away from baseline. Preventive causes in the noisy-OR/AND model are usually understood in this way. A second way of "preventing" is to move *E* towards its lower bound. These *reducers* are the natural opposite of standard generative causes, as they shift *E* downwards while generators shift *E* upwards. The important distinction here is whether the preventer influences the effect directly (i.e., is a reducer), or indirectly through the elimination of other causal influences (i.e., is a blocker).

As a practical example, suppose *Heart Rate* is our effect variable. There are many generative causes that increase heart rate, such as stress or exercise. Beta blockers and other anxiety-reducing medications function as blockers, as they prevent (some of) those generative causes from having any influence while not suppressing *Heart Rate* below its natural baseline (for that individual). In contrast, most anesthetics are reducers of *Heart Rate*, as they actively slow the heart, potentially even below its natural baseline, depending on exactly which causes are active. Of course, if we model *Heart Rate* as simply "low" or "high" (where "low" is the baseline), then

these two different types of drugs will appear indistinguishable. The importance of distinguishing reducers from blockers becomes apparent only when we move to situations in which the lower bound and baseline values need not coincide.

Before turning to the fully general mathematical framework for causal capacities, we must address a potential ambiguity about a capacity's "causal strength" w_i. The standard interpretation in the noisy-OR/AND model is that w_i expresses the probability that the capacity is "active," where an active cause deterministically produces the effect (unless a suitable blocker is also active). This interpretation is inappropriate when causes are more than binary, as "probability of activation" neglects the (presumed) importance of the magnitude of the cause variable.[7] Instead, we will understand w_i (for generators and reducers) for a capacity C_i to be the expected change in E's value when C_i increases by one unit and every other factor is at its baseline value. That is, w_i is computed by starting in the state in which every causal factor is at baseline, and then determining the *expected* change in E when C increases by one unit.[8] This interpretation implies that w_i depends on C_i's scale, but this should be expected given the predictive function of causal strengths. Notice that, if all causes and the effect are binary, then the expected change and probability of activation interpretations of w_i are mathematically identical. The expected change

[7] We can retain the "probability of activation" interpretation if the effect is the only many-valued variable, in which case the natural representations are noisy-ADD or noisy-MAX functions (Heckerman and Breese 1996).

[8] If causal strength depends on C's value, then the choice to measure from C's baseline is potentially a substantive one. However, since we assume causes have independent monotonic influences, we can always transform the scale for C so that E is a linear function of C's value.

interpretation, however, also naturally applies to systems in which some factors can take on more-than-two values, and so I use it in the next section.

3. A General, Privileged Mathematical Representation

Now that we have done the necessary conceptual clarification, we can develop a general, privileged mathematical representation of causal capacities when the causes and effect need not be binary. Throughout, I use lower-case letters to denote the value of a variable; for example, e is the value of the effect E. Without loss of generality, we can assume E's baseline value is zero and $e \in [-L, U]$, where at least one of L, U is greater than zero (else E is always zero). Note that the baseline can be the same as the lower bound ($L = 0$, $U > 0$); same as the upper bound ($L > 0$; $U = 0$); or a strictly intermediate value (L, $U > 0$). As noted above, three different types of causal capacities must be incorporated into the mathematical framework: generators G_i and reducers R_j that (probabilistically) increase and decrease the value of E, respectively; and blockers B_k that (probabilistically) prevent any other causal capacities from influencing E. For all three types of causes, their values must also be able to range over more than just $\{0, 1\}$. For mathematical convenience, we represent the "inactive" state of each cause by 0, so that the influence on E (when only C is active) is the product of C's magnitude (i.e., its distance from zero) and its causal strength (i.e., the expected change in E given that the cause increased by one unit).[9]

Consider first the case with only generators G_i with values g_i. In this situation, E can only be pushed upwards from the baseline, and so $e \in [0, U]$. The natural mathematical framework

[9] Recall from fn. 8 that the independent causal influences have all been transformed so that they are linear (in C) influences of E, and so this product for a single generator or reducer is always less than the relevant upper or lower bound, respectively.

simply uses normalization to convert this case to (a continuous version of) the noisy-OR model: (i) "normalize" *E* and the causal strengths to the [0, 1] interval; (ii) use the uniquely privileged (Cozman 2004) noisy-OR model; and then (iii) transform the result back to the [0, *U*] interval. The noisy-OR/AND model was defined in equations (1) and (2) in terms of the probability of *E* given its causes, but we can (and should, in the present context) instead regard those equations as providing the expectation of *E*. The natural, privileged mathematical representation for the expectation of *E* in this situation is thus:[10]

$$\mathbb{E}(E) = U[1 - \prod_{i=1}^{g} \frac{U - w_i g_i}{U}]$$

Since reducers are naturally understood as "negative generators," we can model the impact of a set of reducers R_j with values r_j in the same way, though their "normalization" is relative to *L* rather than *U*. The resulting expectation of *E* is simply the difference between these (normalized and combined) influences:

$$\mathbb{E}(E) = U[1 - \prod_{i=1}^{g} \frac{U - w_i g_i}{U}] - L[1 - \prod_{j=1}^{r} \frac{L - w_j r_j}{L}]$$

Finally, blockers B_k with values b_k fill the role of preventers in the noisy-OR/AND model of equation (2): the (probabilistic) activation of their causal capacities prevents the expression of any other causal capacities, and so they act as a probabilistic "switch" on the previous equation. The causal strengths of the blocking capacities are thus best understood as "increase (per unit change in the blocker) in probability of complete blocking." The resulting full mathematical equation is:

[10] I show below that this equation is well-behaved even when $U = +\infty$.

$$\mathbb{E}(E) = \prod_{k=1}^{b}(1 - w_k b_k)\Big[U[1 - \prod_{i=1}^{g}\frac{U - w_i g_i}{U}] - L[1 - \prod_{j=1}^{r}\frac{L - w_j r_j}{L}]\Big] \quad (3)$$

Equation (3) is the natural generalization of the noisy-OR/AND model to cases with many-valued variables and distinct baseline and lower bound for *E*. It thus provides the privileged mathematics of causal capacities for precisely the same reasons as the noisy-OR/AND model for the special case of binary variables. To see that it provides such a generalization, consider the special case that was the focus of the previous section: $L = 0$, $U = 1$, and all of the causal factors are restricted to $\{0, 1\}$. Since *L* is equal to the baseline, there are no "reducing" causal capacities: for any putative reducer *R*, the expected change in *E* from a unit change in *R* (when all other causes are absent) is always zero, and so w_R is always zero. And since the causal factors are restricted to $\{0, 1\}$, the b_k and g_i variable values can be replaced with delta functions. The resulting equation (when we substitute in *U* and *L*) is simply equation (2), the noisy-OR/AND model. That is, the equations and claims of the previous section are all special cases of the generalization provided here.

Equation (3) provides a privileged mathematics for arbitrary variable ranges and causal capacities, in the sense (previously articulated) that it captures the plausible, intuitive features of "standard" cases, and therefore can serve as a natural default representational framework. It is particularly interesting to consider another special case. Suppose $E \in [-\infty, +\infty]$ and (for the moment) that there are no blockers. It is not obvious how to use equation (3) in this situation, since direct substitution of *L* and *U* yields infinities throughout the equation. If we instead

consider the limit of equation (3) as L and U go to infinity, we find that the expectation of E is given by:[11]

$$\mathbb{E}(E) = \sum_{i=1}^{g} w_i g_i - \sum_{j=1}^{r} w_j r_j \tag{4}$$

That is, the natural mathematical equation for (the expectation of) E in this special case is simply a linear function of the causal capacities. Having seen equation (4), it is straightforward to incorporate blockers, as that initial product term will simply act to globally attenuate the linear impact (on the expectation of E) of the generators and reducers.

Equation (3) provides a measure of unification to equations (2) and (4): despite their substantial mathematical differences, both noisy-OR/AND and linear models are special cases of the more general, privileged mathematical characterization of causal capacities. That is, this framework suggests that noisy-OR/AND and linear models have the same conceptual and mathematical basis, and the different models arise simply based on whether the variables are binary or continuous/real-valued. In particular, this unification helps to explain why so many mathematical results that hold for linear models also hold for noisy-OR/AND models, and vice versa. For example, the conditions for model parameter identifiability are essentially the same for noisy-OR/AND models (Hyttinen, Eberhardt, and Hoyer 2011) and linear models (Hyttinen, Eberhardt, and Hoyer 2012). Similarly, we find basically the same conditions and statistical tests

[11] *Proof sketch*: For the generators in equation (3), separate the fraction terms into differences and expand the product to yield: $U\,[1 - (1 - \Sigma\,(w_i g_i / U) + \mathbf{C})] = [\,\Sigma\, w_i g_i - U\mathbf{C}]$, where $\mathbf{C}$ is the rest of the product expansion. Every term in $\mathbf{C}$ has at least U^2 in the denominator, and so as $U \to +\infty$, $U\mathbf{C} \to 0$. Thus, as $U \to +\infty$, we are left with only the sum. The same reasoning yields the sum for reducers.

for discovering an unobserved common cause of multiple observed effects given either a noisy-OR/AND model (Danks and Glymour 2001; Pearl 1988) or a linear model (Spirtes et al. 2000). This overlap in the models' mathematical properties is much less surprising given that they (arguably) derive from a single, more general equation (though their properties are not identical, since the different variable value ranges do sometimes matter).

This mathematical connection can also provide us with insights into human cognition. I earlier noted that the noisy-OR/AND model emerged partly from work in cognitive psychology on one "natural" way that people seem to represent causal strengths in the world, at least when we have binary causes and effects (Cheng 1997; Danks 2014; Holyoak and Cheng 2011). At the same time, there are competing theories of human causal learning—variants of the Rescorla-Wagner model and its long-run counterpart, the conditional ΔP theory (Danks 2003)—in which people represent causal capacities as combining linearly (Danks 2007). Relatedly, there is a long history of psychological research on function approximation that has shown that people find linear functions easier to learn (e.g., McDaniel and Busemeyer 2005; DeLosh, Busemeyer, and McDaniel 1997; and references therein), and even have a significant bias in favor of understanding the world in terms of linear functions (Kalish, Griffiths, and Lewandowsky 2007). Equation (3) provides a measure of theoretical unification for these disparate psychological results: noisy-OR/AND and linearity are not theoretical competitors, but rather different aspects of the same general assumptions or preferences about causal capacities. That is, we need not ask whether noisy-OR/AND *or* linearity is correct, since each is the natural representation *for a particular domain of variable values*.[12]

[12] This observation suggests that people in causal learning experiments might systematically shift between noisy-OR/AND and linearity based solely on the variable value ranges. Unfortunately,

4. Conclusions

The philosophical literature on causal capacities has principally asked metaphysical and epistemological questions, rather than representational ones. At the same time, the psychological and machine learning literature on causal capacities has largely focused on the special case of binary causal factors and a binary effect. By generalizing beyond that special case, we thereby obtain a natural, privileged framework for representing causal capacities that independently influence some effect.[13] Moreover, this generalized framework provides further conceptual clarification about causal capacities, as it reveals distinctions (e.g., between the lower bound and the baseline value) that have previously been relatively little-explored in the psychological and machine learning literatures. This mathematical framework also has significant practical and theoretical impacts, as it provides a natural way to unify disparate equations—in particular, the noisy-OR/AND and linear models—that have previously been viewed (in machine learning and cognitive science) as competitors, or at least independent of one another. The widespread use and value of such models is eminently explainable when we understand them as deriving from the

---

cover stories for those experiments almost never explicitly provide value ranges, and we do not know what participants infer about the possible variable values. Anecdotally, though, this type of switching would explain some otherwise puzzling empirical data.

[13] One open question is whether there are also privileged equations for $P(E)$. As a promising first step, we can prove: if there is one generative cause and the initial $P(E)$ is uniform over $[-L, U]$, then the "update" equation $P_{new}(E = e|G = g) = P_{old}(E = U\frac{e - w_g g}{U - w_g g})$ naturally satisfies all of the desiderata provided throughout the paper (including the desired expectation). It is unknown whether other results of this type can be obtained.

same fundamental framework and equation. This privileged framework provides a precise, formal representation that can significantly constrain and inform our attempts to better understand causal capacities.

References

Cartwright, Nancy. 1989. *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

———. 1999. *The Dappled World: a Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

———. 2002. "Against Modularity, the Causal Markov Condition, and Any Link Between the Two: Comments on Hausman and Woodward." *The British Journal for the Philosophy of Science* 53: 411–53.

———. 2007. *Hunting Causes and Using Them: Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press.

Cheng, Patricia W. 1997. "From Covariation to Causation: a Causal Power Theory." *Psychological Review* 104: 367–405.

Cozman, Fabio G. 2004. "Axiomatizing Noisy-OR." In *Proceedings of the 16th European Conference on Artificial Intelligence*.

Danks, David. 2003. "Equilibria of the Rescorla-Wagner Model." *Journal of Mathematical Psychology* 47: 109–21.

———. 2007. "Causal Learning from Observations and Manipulations." In *Thinking with Data*, edited by Marsha C. Lovett and Priti Shah, 359–388. Mahwah, NJ: Lawrence Erlbaum.

———. 2014. *Unifying the Mind: Cognitive Representations as Graphical Models*. Cambridge, MA: The MIT Press.

Danks, David, and Clark Glymour. 2001. "Linearity Properties of Bayes Nets with Binary Variables." In *Proceedings of the 17th Annual Conference on Uncertainty in Artificial Intelligence*, edited by Jack Breese and Daphne Koller, 98–104. San Francisco: Morgan

Kaufmann.

DeLosh, Edward L., Jerome R. Busemeyer, and Mark A. McDaniel. 1997. “Extrapolation: the Sine Qua Non for Abstraction in Function Learning.” *Journal of Experimental Psychology: Learning, Memory, & Cognition* 23 (4): 968–86.

Glymour, Clark. 1998. “Learning Causes: Psychological Explanations of Causal Explanation.” *Minds and Machines* 8: 39–60.

Good, I. J. 1961. “A Causal Calculus (I).” *British Journal for the Philosophy of Science* 11 (44): 305–18.

Hausman, Daniel M., and James Woodward. 1999. “Independence, Invariance and the Causal Markov Condition.” *The British Journal for the Philosophy of Science* 50: 521–83.

———. 2004. “Modularity and the Causal Markov Assumption: a Restatement.” *The British Journal for the Philosophy of Science* 55 (1): 147–61.

Heckerman, David, and John S. Breese. 1994. “A New Look at Causal Independence.” In *Proceedings of the 10th Annual Conference on Uncertainty in Artificial Intelligence*, 286–92. Morgan Kaufmann.

———. 1996. “Causal Independence for Probability Assessment and Inference Using Bayesian Networks.” *IEEE Transactions on Systems, Man, and Cybernetics: Part a Systems and Humans* 26 (6): 826–31.

Heil, John. 2005. *From an Ontological Point of View*. New York: Oxford University Press.

Holyoak, Keith J., and Patricia W. Cheng. 2011. “Causal Learning and Inference as a Rational Process: the New Synthesis.” *Annual Review of Psychology* 62: 135–63.

Hyttinen, Antti, Frederick Eberhardt, and Patrik O. Hoyer. 2011. “Noisy-or Models with Latent Confounding.” In *Proceedings of the 27th Conference on Uncertainty in Artificial*

Intelligence.

———. 2012. “Learning Linear Cyclic Causal Models with Latent Variables.” *Journal of Machine Learning Research* 13: 3387–3439.

Kalish, Michael L., Thomas L. Griffiths, and Stephan Lewandowsky. 2007. “Iterated Learning: Intergenerational Knowledge Transmission Reveals Inductive Biases.” *Psychonomic Bulletin & Review* 14: 288–294.

Kim, Jin H., and Judea Pearl. 1983. “A Computational Model for Causal and Diagnostic Reasoning in Inference Systems.” In *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, 190–93. San Francisco: Morgan Kaufmann.

Lucas, Peter J. F. 2005. “Bayesian Network Modeling Through Qualitative Patterns.” *Artificial Intelligence* 163: 233–63.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. “Thinking About Mechanisms.” *Philosophy of Science* 67 (1): 1–25.

Martin, C. B. 2008. *The Mind in Nature*. Oxford: Oxford University Press.

McDaniel, Mark A., and Jerome R. Busemeyer. 2005. “The Conceptual Basis of Function Learning and Extrapolation: Comparison of Rule-Based and Associative-Based Models.” *Psychonomic Bulletin & Review* 12 (1): 24–42.

Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible iInference*. San Francisco: Morgan Kaufmann Publishers.

Spirtes, Peter, Clark Glymour, and Richard Scheines. 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: The MIT Press.

Srinivas, Sampath. 1993. “A Generalization of the Noisy-OR Model.” In *Proceedings of the 9th Annual Conference on Uncertainty in Artificial Intelligence*, 208–15.