Chapter 1

A Quantum Bayes Net Approach to Causal Reasoning

Jennifer S. Trueblood*, Percy K. Mistry

Department of Cognitive Science, University of California, Irvine, CA, USA

Emmanuel M. Pothos

Department of Psychology, City University London, London, UK

When individuals have little knowledge about a causal system and must make causal inferences based on vague and imperfect information, their judgments often deviate from the normative prescription of classical probability. Previously, many researchers have dealt with violations of normative rules by elaborating causal Bayesian networks through the inclusion of hidden variables. While these models often provide good accounts of data, the addition of hidden variables is often *post hoc*, included when a Bayes net fails to capture data. Further, Bayes nets with multiple hidden variables are often difficult to test. Rather than elaborating a Bayes net with hidden variables, we generalize the probabilistic rules of these models. The basic idea is that any classic Bayes net can be generalized to a quantum Bayes net by replacing the probabilities in the classic model with probability amplitudes in the quantum model. We discuss several predictions of quantum Bayes nets for human causal reasoning.

1. Background and motivation

Human causal reasoning has intrigued scholars from a number of different fields including philosophy, cognitive science, and developmental psychology. Traditionally, models of causal reasoning have been based on classical probability theory. Early models were centered around the notion that people reason about causes and effects by observing the covariation between events [1, 2]. These models are rooted in ideas from Hume [3], who hy-

*jstruebl@uci.edu

pothesized that our sensory system infers causation through the constant conjunction of events. While covariational models can account for many phenomena, all models in this class have one distinct problem - the covariation of events does not necessarily imply a causal relationship between the events. For example, over the past century the number of pirates has decreased while average global temperatures have increased. While these two events negatively covary, no one would recommend becoming a pirate as a solution to global warming.

Cheng and Novick [4, 5] attempted to overcome these issues by incorporating domain specific prior knowledge with covariational information in their *power PC theory*. In their theory, people infer causal relationships to understand observed regularities in the occurrence of events. The model can explain why covariation sometimes implies causation and sometimes it does not. Despite the many successes of power PC theory, some studies have demonstrated that people's judgments often deviate from the model's predictions [6, 7].

Perhaps the most successful classical approach to modeling human causal reasoning is the one based on causal Bayesian networks (or causal Bayes nets for short). In a causal Bayes net, causal relationships are represented by Bayes' calculus and model predictions are typically accepted as normative [8, 9]. This class of models has been shown to account for causal inferences driven by observation, intervention-based, and counterfactual processes [10], which are often difficult to discriminate with other classical approaches. Causal Bayes nets have also been used to explain how people learn causal relationships over time [11, 12]. Attempts have even been made to combine more traditional approaches, such as power PC theory, with causal Bayes nets [11, 13]. More generally, causal Bayes nets have been applied to classification [14, 15], decision making [16], and structured knowledge [17].

Despite their many successes in accounting for human causal reasoning, recent studies have suggested that people's judgments sometimes deviate from the normative prescription of causal Bayes nets. In particular, people commonly violate a condition of causal Bayes nets called the local Markov property. This property is fundamental to causal Bayes nets and, in part, defines the probabilisitic rules of the net. The property states that if we have knowledge about all of the possible causes of some event A, then the descendants (i.e., the effects) of A may tell us information about A, but the non-descendants (i.e., noneffects) do not. Recently, several studies have shown the people often violate the local Markov condition [18–23]. Relatedly, Fernbach et al. [24] found the people often ignore relevant variables in their causal judgments. In their study, people often ignored alternative causes in predictive reasoning situations (i.e., reasoning about an effect given the causes), but not in diagnostic reasoning situations (i.e., reasoning about the causes given the effects).

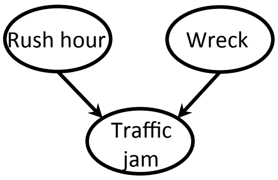
One way to overcome the issues mentioned above is to extend a causal Bayes net through the inclusion of hidden variables. These latent variables are considered part of the mental reconstruction of a causal system by an individual, but are never explicitly described in the experiment. Such approaches often provide good accounts of data [20]; however, they are difficult to conclusively test because the variables are never mentioned by the experimenter. Further, hidden variables are typically added to a causal Bayes net in a *post hoc* manner, after a basic causal Bayes net fails to explain behavioral phenomena. As an alternative approach, we propose generalizing causal Bayes net by altering the probabilistic rules that govern the net. The basic ideas is to replace the classical probabilities in a causal Bayes net with quantum ones, thereby producing quantum Bayes nets [25, 26]. In the next sections, we review causal Bayes nets in more detail and introduce quantum Bayes nets.

2. Causal Bayes nets

A causal Bayes net represents casual relationships between variables through a directed acyclic graph (DAG), which describes a set of random variables and their conditional dependencies. For example, suppose you are driving to work and run into a traffic jam. You might consider two possible causes for the traffic jam - (1) normal rush hour traffic or (2) an automobile wreck. In this example, the three variables - traffic jam, rush hour, and wreck - are each represented as nodes in the DAG (see Figure 1). The edges connecting the nodes represent conditional dependencies. In Figure 1, there is an edge between rush hour and traffic jam as well as wreck and traffic jam. If two nodes are not connected by an edge, then they are conditionally independent. In the example, there is no edge between wreck and rush hour because they are assumed to be conditionally independent.

The probability of a node having a specific value is determined by a probability function that uses the values of parent nodes as inputs. These probabilities are defined in conditional probability tables. Consider the traffic jam scenario where the three variables have two possible values: J = traffic jam is present (true/false), R = rush hour traffic is present

Rush hour	
t	f
.7	.3



Wreck	
t	f
.1	.9

		Traffic jam	
Rush hour	Wreck	t	f
t	t	.9	.1
t	f	.8	.2
f	t	.8	.2
f	f	.5	.5

Fig. 1. A causal Bays net of the traffic jam scenario. There are two possible causes of a traffic jam - rush hour traffic or a wreck. The three binary variables are represented as a DAG with conditional probability tables.

(true/false), and W = wreck is present (true/false). The causal Bayes net can address such questions as "What is the probability that there is a wreck, given that there is traffic jam?" by using the formula for conditional probability:

$$p(W = t|J = t) = \frac{p(J = t, W = t)}{p(J = t)} = \frac{\sum_{j\in\{t,f\}} p(J = t, W = t, R = j)}{\sum_{i,j\in\{t,f\}} p(J = t, W = i, R = j)} \tag{1}$$

where the joint probability function $p(J = t, W = i, R = j) = p(J = t|W = i, R = j)p(W = i)p(R = j)$ because W and R are conditionally independent. The desired probability $p(W = t|J = t)$ is then calculated using the conditional probability tables in Figure 1:

$$p(W = t|J = t) = \frac{(.9 \times .1 \times .7) + (.8 \times .1 \times .3)}{(.9 \times .1 \times .7) + (.8 \times .1 \times .3) + (.8 \times .9 \times .7) + (.5 \times .9 \times .3)} \approx .120 \tag{2}$$

The probabilities for other questions (e.g., $p(W = f|J = f)$) follow similar calculations.

A causal Bayes net obeys the local Markov property, which states that any node in a Bayes net is conditionally independent of its non-descendants (i.e., noneffects) given information about its parents (i.e., direct causes). Consider a causal situation where a variable A causes B and C (represented by the DAG: $B \leftarrow A \rightarrow C$). The local Markov property states that if you know A, then B provides no new information about the value of C. As a result, we have $p(C|A) = p(C|A, B)$. Recently, researchers have found empirical evidence that people's causal judgments do not always obey the local Markov property. In particular, Rehder [20] gave participants causal scenarios with three binary variables (e.g., an economic situation with variables described as trade deficits, interest rates, and retirement savings) and asked them to judge the value of an unknown target variable given information about either or both of the remaining variables. Rehder's [20] results showed that information about non-descendants (i.e., noneffects) influenced judgments even when the values of direct causes were known, thus demonstrating a violation of the local Markov property. (We provide more details about this experiment in a section 4.1.)

Rehder [20] accounted for the observed violations of the local Markov property by elaborating a basic causal Bayes net with an additional variable that was either a shared mediator, shared disabler, or shared cause. For example, a common cause structure where A causes B and C can be augmented in several different ways by the inclusion of a fourth variable D as illustrated in Figure 2. Even though this approach can provide a reasonable account of human judgments, it is difficult to conclusively test because participants are never queried about the latent variable D. As an alternative to using hidden variables, we propose generalizing causal Bayes nets to quantum Bayes nets.

3. Quantum Bayes nets

Following Tucci [25, 27] and Moreira et al. [28] our quantum Bayes nets are constructed by replacing the classical probabilities in the conditional probability tables of a causal Bayes net with quantum probability amplitudes. Consider a simple scenario where there are two causally related, binary (true/false) variables A and B such that $A \rightarrow B$. In quantum probability

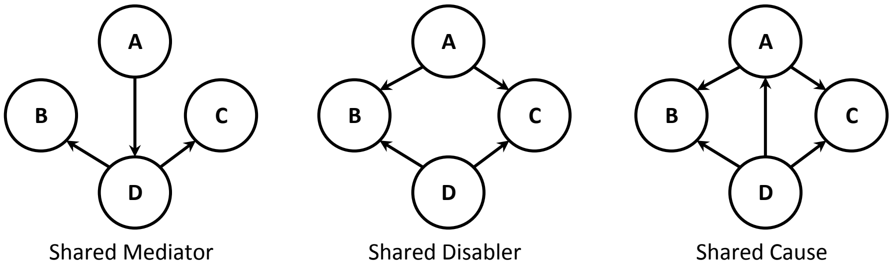


Fig. 2. Three ways to elaborate a common cause structure with an new variable D.

theory, these variables are represented by Hermitian operators:

$$A = a_t P_{a_t} + a_f P_{a_f} \tag{3a}$$

$$B = b_t Q_{b_t} + b_f Q_{b_f} \tag{3b}$$

where a_i and b_i are eigenvalues and P_{a_i} and Q_{b_i} are projectors onto the corresponding eigen-subspaces. The probability of a specific value, such as a_t (the notation a_t is shorthand for $A = t$), is given by Born's rule:

$$p(A = t) = \langle P_{a_t}\psi|\psi\rangle = ||P_{a_t}\psi||^2 \tag{4}$$

where ψ is a unit length state vector representing an individual's knowledge about the different variables.

Now, suppose we want to determine the probability that B is false given that A is true. To answer this question, we first calculate the conditional state vector ψ_{a_t} and then apply Born's rule:

$$p(B = f|A = t) = \langle Q_{b_f}\psi_{a_t}|\psi_{a_t}\rangle = ||Q_{b_f}\psi_{a_t}||^2. \tag{5}$$

These quantum conditional probabilities are then used to define the conditional probability tables associated with the net.

Consider the traffic jam scenario again. With a causal Bayes net, we can answer questions such as "What is the probability that there is a wreck, given that there is a traffic jam?" by calculating joint probabilities such as $p(J = t, W = i, R = j)$. These joint probabilities are determined from the conditional probability tables of the causal Bayes net by writing $p(J = t, W = i, R = j) = p(J = t|W = i, R = j)p(W = i)p(R = j)$. In our quantum Bayes net, we take a similar approach. First, let the three variables traffic jam, wreck, and rush hour be associated with projectors P, Q, and S respectively. We define joint probabilities by Born's rule:

$$p(J = t, W = i, R = j) = ||P_{j_t}\psi_{w_i,r_j}||^2||Q_{w_i}\psi||^2||S_{r_j}\psi||^2 \tag{6}$$

where the conditional state is given by

$$\psi_{w_i,r_j} = \frac{S_{r_j} Q_{w_i} \psi}{||S_{r_j} Q_{w_i} \psi||}. \tag{7}$$

If the projectors Q and S do not commute, then the conditional state will depend on the order in which these two variables are considered so that $\psi_{w_i,r_j} \neq \psi_{r_j,w_i}$. As a result, $p(J = t|W = i, R = j) \neq p(J = t|R = j, W = i)$.

Figure 3 shows a quantum Bayes net generalization of the traffic jam causal Bayes net shown in Figure 1. The probabilities in the causal Bayes net have been replaced by probability amplitudes in the quantum Bayes net. These amplitudes can be related to classical probabilities by squaring the magnitude of the amplitudes. For example, the probability that there was a wreck is given by

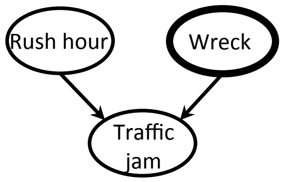
$$\begin{aligned} p(W = t) &= ||.3162e^{i\theta wt}||^2 = (.3162e^{i\theta wt})\overline{(.3162e^{i\theta wt})} \\ &= (.3162e^{i\theta wt})(.3162e^{-i\theta wt}) \\ &= (.3162)^2 e^{i(\theta wt - \theta wt)} = .1 \end{aligned} \tag{8}$$

which is identical to the classical probability in the causal Bayes net. Note that the term $e^{i\theta wt}$ is simply the phase of the amplitude.

When calculating the conditional probabilities of the traffic jam given information about a wreck and rush hour, the order in which the variables wreck and rush hour are considered is important. In the quantum Bayes net shown in Figure 3, we assume that wrecks are always considered before rush hour. Psychologically, this means that an individual thinks about wrecks and rush hour separately, always starting with wrecks. This is represented in the figure by the thick border on the wreck node. If we want to switch the order and have rush hour processed before wrecks, then we would need to define a different set of conditional probabilities. In other words, there are two different conditional probability tables describing the probability of the traffic jam given information about wrecks and rush hour - one table describing the scenario where wrecks is processed before rush hour (shown in Figure 3) and another table describing the scenario where rush hour is considered before wrecks (not shown). The two conditional probability tables are related to one another in a specific manner. In quantum probability theory, noncommutative events (such as wreck and rush hour) are related by a unitary transformation, which preserves lengths and inner products.

For the traffic jam scenario, we started with a causal Bayes net and generalized this to a quantum Bayes net by designating a processing order (wreck before rush hour) and replacing classical probabilities by probability amplitudes. Note that the decision that wrecks should be processed before rush hour was arbitrary. We could have easily chosen the reverse order (rush hour before wrecks). That is, there are at least two different ways to generalize the causal Bayes net in this example. In general, there will often be multiple ways to turn a causal Bayes net into a quantum one. That is, if you start with a well-defined causal Bayes net, there are several ways to convert it into a quantum one because of the arbitrariness in the order of events. This is because noncommutative events result in different conditional probability tables for the same causal scenario. Thus, the behavior of a quantum Bayes net will often be fundamentally different than that of a causal Bayes net.

Rush hour	
t	f
$.8367e^{i\theta rt}$	$.5477e^{i\theta rf}$



Wreck	
t	f
$.3162e^{i\theta wt}$	$.9487e^{i\theta wf}$

		Traffic jam	
Rush hour	Wreck	t	f
t	t	$.9487e^{i\theta jt\|wt,rt}$	$.3162e^{i\theta jf\|wt,rt}$
t	f	$.8944e^{i\theta jt\|wt,rf}$	$.4472e^{i\theta jf\|wt,rf}$
f	t	$.8944e^{i\theta jt\|wf,rt}$	$.4472e^{i\theta jf\|wf,rt}$
f	f	$.7071e^{i\theta jt\|wf,rf}$	$.7071e^{i\theta jf\|wf,rf}$

Fig. 3. A quantum Bayes net generalization of the traffic jam scenario. The wreck node in the DAG has a thick border to indict that it is considered before rush hour. The tables contain probability amplitudes rather than probabilities. These amplitudes were determined from the causal Bayes net shown in Figure 1. The term $e^{i\theta n}$ is the phase of the amplitude associated with situation n. For example, the probability amplitude for $J = t$ given that $W = t$ and $R = t$ (first row of the bottom conditional probability table) has a phase of $e^{i\theta n}$ where the index is $n = jt|wt, rt$.

4. Predictions from quantum Bayes nets

Quantum Bayes nets with noncommutative events generate several interesting predictions regarding human causal judgments. In the following sections, we discuss these predictions and the supporting empirical evidence.

4.1. *Violations of the local Markov condition*

The local Markov property states that any node in a DAG is conditionally independent of its nondescendants when its parents (direct causes) are known. For example, in the common effect structure $A \to C \leftarrow B$, the two causes A and B are conditionally independent. That is, if A and B are binary variables, then $p(A = i|B = t) = p(A = i|B = f)$ for $i \in \{t, f\}$ and likewise when A and B are swapped. In a quantum Bayes net, if A and B do not commute, then there is a dependency between these two variables such that knowing information about B influences our beliefs about A. This dependency leads to violations of the local Markov condition because it implies $p(A = i|B = t) \neq p(A = i|B = f)$. By the definition of conditional probability, $p(A = i|B = j) = p(A = i, B = j)/p(B = j)$. In a causal Bayes net, we have $p(A = i, B = j) = p(A = i)p(B = j)$ because A and B are independent. As a result, $p(A = i|B = j) = p(A = i)$ for all i, j. In a quantum Bayes net, $p(A = i, B = j) = ||P_{a_i} Q_{b_j} \psi||^2$, and if A and B do not commute, then clearly $||P_{a_i} Q_{b_j} \psi||^2 \neq ||P_{a_i} \psi||^2 ||Q_{b_j} \psi||^2$, leading to violations of the local Markov condition.

Rehder [20] showed that people's causal judgments often violate the local Markov condition. In his experiment, participants read information about two causal scenarios involving an unknown target variable and were asked to choose the scenario where the target variable was more probable. For example, in the common effect situation where $A \to C \leftarrow B$, participants had to choose whether the target variable B was more likely to be true in a scenario where $A = t$ or in a scenario where $A = f$. A causal Bayes net predicts that participants' choice proportions should be equal on average because $p(B = t|A = t) = p(B = t|A = f)$ by the local Markov property. However, Rehder [20] found that people tended to select the causal scenario where $A = t$ more often than the alternative where $A = f$. These results suggest that people judged $p(B = t|A = t) > p(B = t|A = f)$. Rehder found similar violations of the local Markov property with other causal structures such as common cause structures ($B \leftarrow A \to C$) and chain structures ($A \to B \to C$).

4.2. *Anti-discounting behavior*

Quantum Bayes nets with noncommutative events can also explain anti-discounting behavior in causal reasoning. The term discounting is used to describe a situation where one cause casts doubts on another cause. For example, in the common effect structure $A \to C \leftarrow B$, knowledge of A could cast doubt on the value of B such that $p(B|C, A) < p(B|C)$. In many causal situations discounting is normative [29]. To see why this might be the case, consider the conditional probabilities $p(B = t|C = t, A = t)$ and $p(B = t|C = t)$. It is normatively correct to judge $p(B = t|C = t) > p(B = t|C = t, A = t)$ because knowing $A = t$ provides sufficient explanation for the value of the effect $C = t$, thus making the other cause B redundant. When the value of A is unknown, there is a greater chance the effect C was brought about by B.

Rehder [20] showed that many people display anti-discounting behavior. In his experiments, people judged an unknown target variable B as highly likely when an alternative cause $A = t$ was known, leading to judgments where $p(B = t|C = t) < p(B = t|C = t, A = t)$. Similar to violations of the local Markov property, quantum Bayes nets explain anti-discounting behavior by the noncommutativity of A and B. When two events do not commute, there is a causal dependency between them. Thus knowing the value of one cause can increase the probability of another cause having a similar value, resulting in judgments that are opposite to those predicted by discounting.

4.3. *Order effects*

If two events are noncommutative, then quantum Bayes nets naturally predict order effects. In an experiment involving the common effect scenario $A \to C \leftarrow B$, participants might be asked to judge $p(C|A, B)$ where they are given information about A before B. An order effects occurs when an individual's final judgment depends on the order in which the information was presented, so that $p(C|A, B) \neq p(C|B, A)$. Causal Bayes nets have difficulty accounting for order effects because they are based on classical probability theory and thus obey the commutative property. That is, $p(A, B|C) = p(B, A|C)$ implies $p(C|A, B) = p(C|B, A)$ by Bayes' rule. Classical models can be extended to account for order effects, but this involves introducing extra variables such as O_1, the event that A is presented first, and O_2, the event that A is presented second. With these additional events, then it is possible to have $p(C|A, B, O_1) \neq p(C|A, B, O_2)$. However, without a theory about O_1 and O_2, this model simply redescribes the empirical result. Moreover, in most studies of order effects, the presentation order is randomly determined, yielding order information such as O_1 and O_2 irrelevant.

Many empirical studies have documented the importance of order information in human judgments [30]. Order effects have been observed in a number of different domains ranging from judging the probability of selecting balls from urns [31] to judging the guilt of a defendant in a mock trial [32, 33]. Trueblood and Busemeyer [34] demonstrated that order effects also arise in causal reasoning. In their experiment, participants were asked to make judgments about ten different causal scenarios involving a single effect and two binary (present / absent) causes. For example, in one scenario, participants read a description about a college sophomore, Liz, with a 3.0 GPA and were asked to judge the likelihood that Liz will earn an A in her social psychology class. In one condition, participants first read information about how Liz hopes to study social work in graduate school (the present cause) followed by information that Liz does not make any changes to her study habits (the absent cause). In a different condition, participants read the information in the reverse order.

The participants (N = 113) provided likelihood judgments of the effect (e.g., Liz receiving an A in social psychology) on a 0 to 100 scale at three different time points: (1) before reading either cause, (2) after reading one of the causes, and (3) after reading the other cause. Participants judged the present cause before the absent cause in a random half of the scenarios. In the other half of the scenarios, the order of the causes was reversed. The results of the experiment showed a large, significant order effect ($p < 0.001$) across all ten scenarios.

5. Conclusions

Arguably, causal Bayes nets have been one of the most successful modeling approaches to human causal reasoning. These models can account for a range of findings including casual deductive and inductive reasoning across a number of different scenarios. Further, causal Bayes nets have been applied to other domains such as decision-making [16] and classification [14, 15, 35].

Despite their many successes, recent experiments have shown that people's judgments often deviate from the predictions of causal Bayes nets. In particular, people often violate the local Markov property of causal Bayes nets. To overcome such issues, researchers have suggested augmenting causal Bayes nets with additional variables [20]. These hidden variables provide increased model flexiblity, allowing causal Bayes nets to accommodate a wider range of behavior. However, the addition of such variables is often *post hoc* and they are difficult to conclusively test experimentally.

As an alternative, we suggest generalizing causal Bayes nets using quantum probability theory. In our approach, we do not elaborate a causal Bayes net with additional nodes and edges. Rather, we change the probabilistic rules used to perform inference. By using quantum probabilities instead of classical ones, we can allow for variables to be noncommutative. By using noncommutative events in our quantum Bayes nets, we can account for a number of different behavioral phenomena including violations of the local Markov condition, anti-discounting behavior, and order effects.

Quantum probability theory has been successfully applied to a range of different problems in the cognitive and decision sciences including interference effects in perception [36], violations of the sure thing principle [37], violations of dynamic consistency [38], conjunction and disjunction fallacies [39], order effects in survey questions [40], and the interference of categorization on decision-making [41]. We believe that quantum probability theory also has great potential to explain human causal reasoning. The results discussed in this chapter make us optimistic about this approach in the future.

Acknowledgements

JST and PKM were supported by NSF SES-1326275. EMP was supported by Leverhulme Trust grant RPG-2013-00 and Air Force Office of Scientific Research (AFOSR), Air Force Material Command, USAF, grants FA 8655-13-1-3044. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

References

[1] H. M. Jenkins and W. C. Ward, Judgment of contingency between responses and outcomes, *Psychological Monographs: General and Applied.* **79**, 1–17 (1965).
[2] H. H. Kelley, The processes of causal attribution, *American Psychologist.* **28**, 107–128 (1973).
[3] D. Hume, *A treatise of human nature (2nd ed.).* Clarendon Press. (Original work published 1739), Oxford, England (1987).

[4] P. W. Cheng, From covariation to causation: A causal power theory, *Psychological Review.* **104**, 367–405 (1997).
[5] L. R. Novick and P. W. Cheng, Assessing interactive causal influence, *Psychological Review.* **111**(2), 455–485 (2004).
[6] K. Lober and D. R. Shanks, Is causal induction based on causal power? critique of cheng (1997)., *Psychological Review.* **107**(1), 195–212 (2000).
[7] P. A. White, The power pc theory and causal powers: Comment on cheng (1997) and novick and cheng (2004), *Psychological Review.* **112**(3), 675–682 (2005).
[8] J. H. Kim and J. Pearl. A computational model for causal and diagnostic reasoning in inference systems. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 190–193 (1983).
[9] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann (1988).
[10] Y. Hagmayer, S. A. Sloman, D. A. Lagnado, and M. R. Waldmann. Causal reasoning through intervention. In eds. A. Gopnik and L. Schulz, *Causal learning: Psychology, philosophy, and computation*, pp. 86–100. Oxford University Press, Oxford (2007).
[11] T. L. Griffiths and J. B. Tenenbaum, Structure and strength in causal induction, *Cognitive Psychology.* **51**, 334–384 (2005).
[12] T. L. Griffiths and J. B. Tenenbaum, Theory-based causal induction, *Psychological Review.* **116**(4), 661–716 (2009).
[13] H. Lu, A. L. Yuille, M. Liljeholm, P. W. Cheng, and K. J. Holyoak, Bayesian generic priors for causal learning., *Psychological review.* **115**(4), 955 (2008).
[14] B. Rehder and S. Kim, Classification as diagnostic reasoning, *Memory and Cognition.* **37**, 715–729 (2009).
[15] B. Rehder and S. Kim, Causal status and coherence in causal-based categorization, *Journal of Experimental Psychology: Learning, Memory, and Cognition.* **36**, 1171–1206 (2010).
[16] Y. Hagmayer and S. A. Sloman, Decision makers conceive of themselves as interveners, *Journal of Experimental Psychology: General.* **128**, 22–38 (2009).
[17] C. Kemp and J. B. Tenenbaum, Structured statistical models of inductive reasoning, *Psychological Review.* **116**, 20–58 (2009).
[18] B. M. Rottman and R. Hastie, Reasoning about causal relationships: Inferences on causal networks, *Psychological Bulletin.* **140**(1), 109–139 (2014).
[19] J. Park and S. A. Sloman, Mechanistic beliefs determine adherence to the markov property in causal reasoning, *Cognitive Psychology.* **67**(4), 186–216 (2013).
[20] B. Rehder, Independence and dependence in human causal reasoning, *Cognitive Psychology.* **72**, 54–107 (2014).
[21] P. M. Fernbach and S. A. Sloman, Causal learning with local computations, *Journal of Experimental Psychology: Learning, Memory, and Cognition.* **35**, 678–693 (2009).
[22] M. R. Waldmann, P. W. Cheng, Y. Hagmayer, and A. P. Blaisdell. Causal learning in rats and humans: A minimal rational model. In eds. N. Chater

and M. Oaksford, *The Probabilistic Mind. Prospects for Bayesian Cognitive Science*, pp. 453–484. Oxford: University Press, Oxford (2008).

[23] Y. Hagmayer and M. R. Waldmann. A constraint satisfaction model of causal learning and reasoning. In eds. W. D. Gray and C. D. Schunn, *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, pp. 405–410, Erlbaum, Mahwah, NJ (2002).

[24] P. M. Fernbach, A. Darlow, and S. A. Sloman, Neglect of alternative causes in predictive but not diagnostic reasoning, *Psychological Science*. **21 (3)**, 329–336 (2010).

[25] R. R. Tucci, Quantum bayesian nets, *International Journal of Modern Physics, B*. **9**, 295–337 (1995).

[26] J. R. Busemeyer and P. D. Bruza, *Quantum Models of Cognition and Decision*. Cambridge University Press, Cambridge, UK (2012).

[27] R. R. Tucci, An introduction to quantum bayesian networks for mixed states, *arXiv preprint arXiv:1204.1550* (2012).

[28] C. Moreira and A. Wichert, Interference effects in quantum belief networks, *Applied Soft Computing*. **25**, 64–85 (2014).

[29] M. W. Morris and R. P. Larrick, When one cause casts doubt on another: A normative analysis of discounting in causal attribution, *Psychological Review*. **102**(2), 331–355 (1995).

[30] R. M. Hogarth and H. J. Einhorn, Order effects in belief updating: The belief-adjustment model, *Cognitive Psychology*. **24**, 1–55 (1992).

[31] J. C. Shanteau, An additive model for sequential decision making, *Journal of Experimental Psychology*. **85**, 181–191 (1970).

[32] A. Furnham, The robustness of the recency effect: Studies using legal evidence, *Journal of General Psychology*. **113**, 351–357 (1986).

[33] L. Walker, J. Thibaut, and V. Andreoli, Order of presentation at trial, *Yale Law Journal*. **82**, 216–226 (1972).

[34] J. S. Trueblood and J. R. Busemeyer, A quantum probability account of order effects in inference, *Cognitive Science*. **35**, 1518–1552 (2011).

[35] B. Rehder, Categorization as causal reasoning, *Cognitive Science*. **27**, 709–748 (2003).

[36] E. Conte, Y. A. Khrennikov, O. Todarello, A. Federici, and J. P. Zbilut, Mental states follow quantum mechanics during perception and cognition of ambiguous figures, *Open Systems & Information Dynamics*. **16**, 1–17 (2009).

[37] E. M. Pothos and J. R. Busemeyer, A quantum probability explanation for violations of 'rational' decision theory, *Proceedings of the Royal Society B*. **276 (1165)**, 2171–2178 (2009).

[38] J. R. Busemeyer, Z. Wang, and J. S. Trueblood. Hierarchical bayesian estimation of quantum decision model parameters. In ed. J. R. Busemeyer, *QI 2012, LNCS 7620*, Springer-Verlag, Berlin, Germany (2012).

[39] J. R. Busemeyer, E. Pothos, R. Franco, and J. S. Trueblood, A quantum theoretical explanation for probability judgment errors, *Psychological Review*. **118**, 193–218 (2011).

[40] Z. Wang and J. R. Busemeyer, A quantum question order model supported by empirical tests of an a priori and precise prediction, *Topics in Cognitive*

Science. **5**, 689–710 (2013).
[41] J. R. Busemeyer, Z. Wang, and A. Lambert-Mogiliansky, Comparison of markov and quantum models of decision making, *Journal of Mathematical Psychology.* **53**, 423–433 (2009).

Index