

Learning, prediction and causal Bayes nets

Clark Glymour

Carnegie Mellon University, Pittsburgh, and the Institute for Human and Machine Cognition, University of West Florida, Pensacola, USA

Recent research in cognitive and developmental psychology on acquiring and using causal knowledge uses the causal Bayes net formalism, which simultaneously represents hypotheses about causal relations, probability relations, and effects of interventions. The formalism provides new normative standards for reinterpreting experiments on human judgment, offers a precise interpretation of mechanisms, and allows generalizations of existing theories of causal learning. Combined with hypotheses about learning algorithms, the formalism makes predictions about inferences in many experimental designs beyond the classical, Pavlovian cue \rightarrow effect design.

Understanding the causal structure of the world is a fundamental human capacity that allows human beings to control and predict their physical and social environments. Both theoretical and experimental work on causal judgment have chiefly relied on conceptions first formulated in the 19th and early 20th centuries: Boole [1] and Frege's [2] construal of causal claims as logical conditionals, Pavlov's [3] classical conditioning, and Piaget's [4] account of developmental stages. With the exception of studies of perceptual clues for causal judgments, [5–7], late 20th century cognitive psychology chiefly added some computational mechanisms: hybrid computational models for the logical conception [8], and analogies with linear neural net learners for classical conditioning, [9,10], while developmental psychology has had occasional recourse to associationist mechanism related to classical conditioning [11]. A Pavlovian cue \rightarrow effect design remained common in experimental studies of adult causal judgment: 'cues' or potential causes preceded effects, and as, or after, subjects observed sequences of cases of cues and effects (or their absences), their judgments of the causal strengths or efficacies of the various cues were elicited.

Causal complexities

Causal relations can be structurally and epistemically more complex than the Pavlovian cue \rightarrow effect design allows. Alternatives include chains of causes, multiple causes influencing one another; interactive causes, unobserved factors influencing both effects and observed potential causes, absence of prior knowledge or time order separating cause and effect, deterministic and probabilistic dependences, interventions that vary some factors while holding others constant, and uses of evidence involving both passive observation and interventions.

Until very recently few of these complexities were allowed in experimental investigations of causal judgment.

All that has begun to change. Psychological studies of how causal understanding develops and is exercised are currently undergoing a revision that cuts across developmental and cognitive psychology, and has undeveloped implications for social psychology and the study of human factors. The common framework of these innovations is the theory of 'causal Bayes nets' that has emerged since 1980 from converging work in statistics, philosophy and artificial intelligence [12–16].

Causal bayes nets

Representation of causal and probability relations

Causal Bayes nets represent causal hypotheses as directed graphs. The Bayes net formalism provides a general connection between causal structure and probability, the Markov assumption, which says that a variable A in a causal graph is independent of all other variables that are not its effects, conditional on its direct causes in the graph – the variables with edges directed into A (see Box 1).

Representation of interventions

The Bayes net formalism represents the effects of interventions on a variable B in a system by introducing a new variable, I. This variable represents the intervention and has a directed edge into the directly manipulated variable, B. When the intervention variable has the value *off*, the system has the original structure and probabilities. However, other values of the intervention variable fix a value (or new probability distribution) for B, break the other edges directed into B, but leave the original probability distribution (conditional on the fixed value of B) intact (Box 2).

Learning algorithms for causal bayes nets

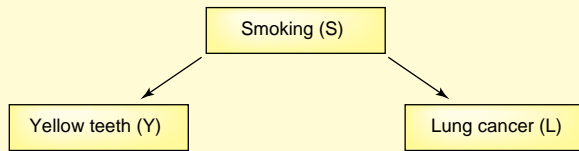
A variety of learning algorithms have been proposed for learning causal Bayes nets from observations and interventions, with or without relevant background knowledge, and they have found many scientific applications, [16–18]). They include algorithms that approximate Bayesian learning, [19], as well as algorithms that identify conditional independence relations and use them to construct a set of possible causal explanations [14] and many combinations and variations of these approaches. Algorithms have also been developed for estimating when the effect of an intervention can be computed from partial causal knowledge, and for computing it (Box 3).

Overview of implications for psychology

The causal Bayes net formalism corrects normative misinterpretations of some experiments with cue \rightarrow effect

Corresponding author: Clark Glymour (cg09@andrew.cmu.edu).

Box 1. Bayes nets, the Markov assumption and conditional independence



TRENDS in Cognitive Sciences

The graph above represents the claim that smoking is a cause of yellowed teeth and lung cancer, but that lung cancer does not cause yellowed teeth and yellowed teeth do not cause lung cancer. It also represents claims about the conditional probability relations among the three variables: for all values of Y, S and L (for example, all combinations of present or absent)

$$\Pr(Y, S, L) = \Pr(Y|L, S) \cdot \Pr(L|S) \cdot \Pr(S) = \Pr(Y|S) \cdot \Pr(L|S) \cdot \Pr(S)$$

where $\Pr(Y = \text{present} | L = \text{absent}, S = \text{present})$, for example, represents the probability of yellowed teeth among smokers without lung cancer. The first equality is necessarily true, but the second is an assumption, the Markov factorization, which says that the joint distribution of all variables is equal to a product of the conditional distributions of each variable on its parents in the graph. The Markov factorization is equivalent, in this example, to the claim that $\Pr(Y|S, L) = \Pr(Y|S)$.

designs, provides generalizations of proposed theories, makes precise notions of mechanism sometimes used in psychological accounts of causal reasoning [20,21] and suggests experiments to distinguish among alternative theories. Supplemented with hypotheses about learning mechanisms, the formalism provides predictions about a range of experiments outside cue \rightarrow effect designs. In what follows I will describe some of this new work, omitting important work on categorization [22,23] which is beyond the scope of this article.

Causal bayes nets and cue \rightarrow effect experimental designs

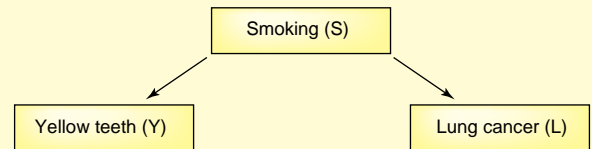
Normative judgments in 'overshadowing'

Baker *et al.* [24] presented subjects with a video set-up in which a tank moved through a minefield toward safety. Subjects could camouflage the tank by pushing a joystick. A plane would sometimes appear in the midst of the tank's traverse. Contingencies were arranged so that the tank reached safety when and only when the plane appeared, and $\sim 65\%$ of the time when the tank was camouflaged. Subjects' judgments of the 'efficacy' of camouflage and of the plane in causing the tank to reach safety were elicited before any trials, after 20 trials and after 40 trials on a -100 to $+100$ scale. Initially, subjects judged the camouflage to be more effective than the plane, but after 40 trials they typically judged the camouflage to have essentially no efficacy and the plane to be very effective. Appealing to Allan's [25] claim that normatively, the efficacy of A to produce B should be judged to be $\Pr(B|A) - \Pr(B|\sim A)$, Baker *et al.* claimed their subjects judgments were non-normative,

<http://tics.trends.com>

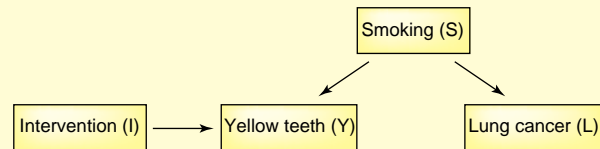
Box 2. Graphical representations of interventions

Starting with the causal system represented by the directed structure:



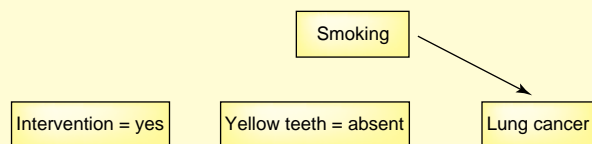
TRENDS in Cognitive Sciences

with the joint probability distribution $\Pr(Y, S, L) = \Pr(Y|S) \cdot \Pr(L|S) \cdot \Pr(S)$, we imagine an intervention that forces everyone to brush their teeth daily with stain removing paste. We reconceive the system above with an expanded structure:



TRENDS in Cognitive Sciences

and probability distribution $\Pr(Y, S, L, I) = \Pr(Y|S, I) \cdot \Pr(L|S) \cdot \Pr(S)$ with the understanding that $\Pr(Y, S, L, I = \text{no intervention}) = \Pr(Y|S) \cdot \Pr(L|S) \cdot \Pr(S)$, the original distribution, and $\Pr(Y, S, L, I = \text{intervention}) = \Pr(Y|I = \text{intervention}) \cdot \Pr(L|S) \cdot \Pr(S) = 0$ unless $Y = \text{present}$, and equals $\Pr(L|S) \cdot \Pr(S)$ otherwise. The act of intervention fixes the value of Y to present, and thus makes Y independent of S. The system *after* the intervention can be more simply represented by the graph:



TRENDS in Cognitive Sciences

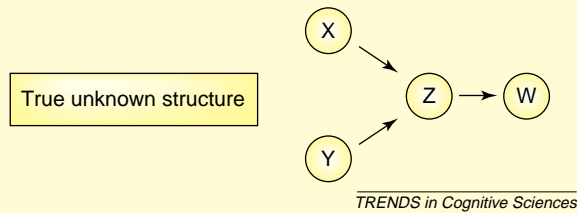
but did accord with simulations of the Rescorla–Wagner model of classical conditioning (Box 4).

But the contingencies and cover story allow a quite different conclusion using the Bayes net framework [26]: the experimental set-up made the appearance of the plane statistically dependent on using the joystick to camouflage the tank, and made reaching safety independent of camouflage conditional on the plane's appearance. Normatively, subjects could have had the causal model shown in Fig. 1, and their judgments after 40 trials accord with it.

Cheng models

Cheng [27] reported a range of experiments on adult subjects using a cue \rightarrow effect design. She found, for example, that in experiments in which subjects think

Box 3. Comparing Bayesian learning and constraint-based learning of Bayes nets



Bayesian learning

- (1) Prior probability distribution $\Pr(G; \theta)$ over all directed acyclic graphs G and probability distributions θ on the variables (vertices in G), with a Markov factorization for G .
- (2) Likelihood function $L(D; G, \theta)$ giving the probability of the observations D conditional on the truth of G, θ .
- (3) Compute the probability of any graph G conditional on the data by using Bayes Theorem and integrating over θ

$$\Pr(G|D) = \frac{\int \Pr(G; \theta)L(D; G, \theta)d\theta}{\Pr(D)}$$

- (4) Find the graphs G such that for all other graphs G^* , $\Pr(G|D) \geq \Pr(G^*|D)$

Constraint-based learning

- (1) Form the complete undirected graph, U
- (2) Estimate from the data which pairs of variables are independent and remove the corresponding edges to form $U1$.
- (3) Estimate from the data which pairs of variables in $U1$ are independent conditional on one of their adjacent neighbors and remove their edges, forming $U2$; continue to form $U3, U4, \dots$ until no more edges are removed.
- (4) Orient $X - Z - Y$ as $X \rightarrow Z \leftarrow Y$ if Z was not conditioned on when removing the $X - Y$ edge.
- (5) Orient remaining undirected edges so as to avoid creating colliders: $\rightarrow V \leftarrow$

there may be unobserved causes of the effect, as well as observed causes, they prefer to suspend judgment about the efficacy of potential observed generative causes when the effect occurs on all trials, and, when the effect never occurs, they prefer to suspend judgment about the efficacy of potential preventive causes. The quantity, $\Pr(C \text{ causes } E \text{ when } C \text{ occurs})$ is the generative causal power of E . She showed that, assuming various independencies, the

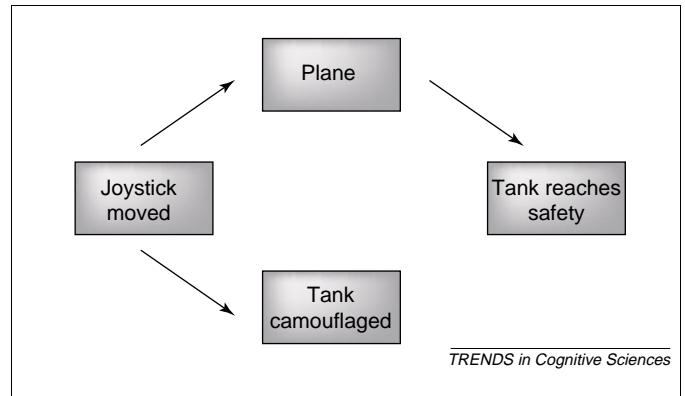


Fig. 1. Normative casual model in making judgments (see text for details).

Box 5. From noisy-or gates to Cheng models

In a noisy-or gate an effect E is assumed to be a Boolean function of its potential causes A, U , and parameters:

$$E = qa A \oplus qu U \tag{1}$$

where qa, qu, E and U all take values $\{0,1\}$, and \oplus is Boolean sum (=1 if and only if either argument is 1). So, letting $A, \text{ etc.},$ stand for $A = 1, \text{ etc.},$ and $\sim A, \text{ etc.},$ stand for $A = 0, \text{ etc.},$ (1) implies:

$$\Pr(E) = \text{pr}(qa A \oplus qu U) = \Pr(qa \cdot A) + \Pr(qu \cdot U) - \Pr(qa \cdot A) \cdot \Pr(qu \cdot U) \tag{2}$$

Therefore $\Pr(E|A, \sim U) = \Pr(qa)$, Cheng's causal power of A to generate E .

If A, U, qa and qu are all independent in probability, (2) becomes:

$$\Pr(E) = \Pr(qa) \cdot \Pr(A) + \Pr(qu) \cdot \Pr(U) - \Pr(qa) \cdot \Pr(A) \cdot \Pr(qu) \cdot \Pr(U) \tag{3}$$

and (3) implies, with some algebra:

$$\Pr(qa) = [\Pr(E|A, \sim U) - \Pr(E|\sim A, \sim U)]/[1 - \Pr(E|\sim A)] \tag{4}$$

which is Cheng's formula for estimating the generative causal power of A when U is thought to be independent of A .

generative causal power of an observed cause can be estimated from observed frequencies. A similar result is shown for preventive causal powers. Such estimates require that the subject select a set of cases – focal sets – in which the occurrence of the cause whose power is to be estimated is judged to be independent of the occurrences of other potential causes. Cheng and Novick [28] have generalized the theory for interactive causes (Box 5).

Cheng's models are known in the Bayes net literature as noisy-or gates (generative) and noisy-and gates (preventive). Cheng's causal powers are probabilities of parameters that specify the occurrence of the effect given

Box 4. Rescorla–Wagner learning

The Rescorla–Wagner procedure estimates that the associative strength of potential cause C_i with the effect, E , after trial $t + 1$ is $V_i = V_i + \Delta V_i$, where ΔV_i is given by:

$$\Delta V_i^t = \begin{cases} 0, & \text{if the cause, } C_i \text{ does not appear in case } t; \\ \alpha_i \beta_1 \left(\lambda - \sum_{\text{Cause } C_j \text{ appears in case } t} V_j \right), & \text{if both } C_i \text{ and } E \text{ appear in case } t; \\ \alpha_i \beta_2 \left(0 - \sum_{\text{Cause } C_j \text{ appears in case } t} V_j \right), & \text{if } C_i \text{ appears and } E \text{ does not in case } t. \end{cases}$$

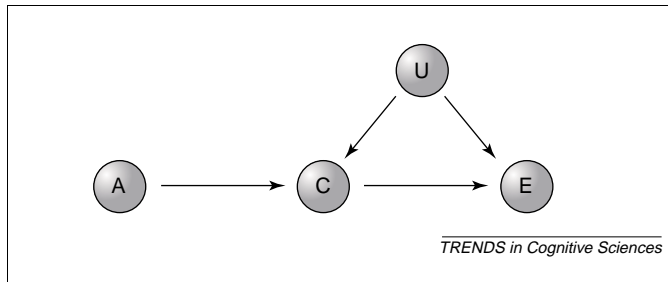


Fig. 2. Graph of an (extended) Cheng model for which the generative causal power of C can be estimated from observations of A, C and E without a focal set.

values (occurrence or non-occurrence) of its causes. When chained together, so that there are sequences of causes, such models automatically satisfy the Markov assumption, and her theory naturally yields a general parameterization of *any* directed acyclic graph, which in turn suggests a variety of as yet untested hypotheses [26]. For example, the generative causal power of C to produce E can sometimes be estimated when it is known that there is an unobserved common cause U of C and E, as in Fig. 2, although there is no focal set.

Using Bayesian learning procedures for networks, Tenenbaum and Griffiths [29] gave an alternative explanation of Cheng's and other data on causal learning in cue \rightarrow effect designs. Waldmann and colleagues, [30], suggested that the Bayes net formalism may be used to represent causal prior knowledge and that such knowledge influences new causal judgments, a phenomenon demonstrated in an intricate set of experiments by Cheng and Lien [31] for prediction of ambiguous cases of subordinate categories from learned causal relations involving superordinate categories.

Beyond the cue \rightarrow effect design: hidden common causes, causal chains, no prior separation of cause and effect, and interventions

Hidden common causes

With a suitable cover story, Danks and Mackenzie had subjects observe values of variables from two causal Bayes nets: $C2 \leftarrow C1 \rightarrow E$ and $C1 \rightarrow C2 \leftarrow U \rightarrow E$. Subjects were not informed about U and U was unobserved. Subjects were asked to judge if C1 is a cause of E and if C2 is a cause of E. A second experiment replaced data from the first structure with data from $C1 \rightarrow E \leftarrow C2$ and a distinct cover story. Danks [32] showed that for the probabilities used in these experiments the Rescorla–Wagner dynamical model yields equilibrium values of associative strength (i.e. values for which the expected change in association on further trials is zero) of a potential cause C1, for example, equal to conditional ΔP , that is, $\Pr(E|C1, \sim C2) - \Pr(E|\sim C1, \sim C2)$. There is no Cheng focal set for this problem. Subjects observed cases until they were prepared to judge whether each observed potential cause was an actual cause of E. There are 256 distinct possible patterns of response. Rescorla–Wagner answers are (yes, no), (yes, yes), (yes, yes), (yes, yes). The Bayes net answers to the four questions are (yes, no), (no, no), (yes, yes), and (no, no). A plurality of subjects identified no causes, and arguably should have been excluded by a pre-test. Of the remainder, 33% of subjects, in the first experiment and 20% of subjects

in the second experiment gave the correct (Bayes net) responses. Only 2 of these 114 subjects gave the conditional ΔP responses.

Causal chains

Lagnado and Sloman [33] gave subjects data on two potential causes of an effect, generated from a Bayes net with structure $A \rightarrow B \rightarrow E$, and required subjects to determine the causal relations among the three variables. Subjects were poor at detecting the correct causal chain, but the design was possibly flawed by a deterministic dependence of B on one value of A and by a bias created by the content of the cover story.

No prior separation of cause and effect

Streyvers and his collaborators gave subjects repeated blocks of 8 trials. All data within a block were determined by one of two alternative structures, $A \rightarrow B \leftarrow C$ or $A \leftarrow B \rightarrow C$, and subjects were given a forced choice, with feedback after each block. Using a Bayesian model of inference, they found groups of subjects that guessed at random, groups that made appropriate Bayesian inferences based only on the last trial in each block, and groups that learned appropriately from the trials within each block. In combination with the results of the Danks–Mackenzie experiments, this work suggests wide individual differences in sophistication of learning strategies.

Interventions

Streyvers and his collaborators also gave adult subjects 18 possible causal graphs on 3 variables, and asked them to choose a best hypothesis from data. Once a hypothesis was chosen, subjects were asked to choose an experimental intervention to test or modify their hypothesis. Subjects showed a bias for manipulating variables hypothesized to be causal sources rather than intermediate variables. Subjects were good at distinguishing common effect structures from others (in agreement with Danks' and Mackenzie's results described above), but were poor at distinguishing chain or common cause structures differing in the middle variable. The accuracy of estimates of structure given by a majority of subjects improved after they acquired data on interventions; a small fraction of subjects made worse estimates.

Markov principles and interventions in developmental psychology

A series of experiments by Gopnik and her collaborators, [34,35], have found that young children (3–4 years) make causal judgments in simple observational cases in accord with the Markov assumption, use co-occurrence information to override spatial contiguity; use causal judgments obtained from observation of co-occurrences to intervene correctly to stop a causal process, and correctly infer causal relations from relatively complex combinations of observed co-occurrences and interventions, distinguishing common effects, common causes and causal chains. For example, observing puppets in correlated motion and a cover story that implies that one of the puppets causes the other, but not which, and an intervention with correlated stopping of puppet A but not of

puppet B, the majority of 4-year-olds correctly infer that the motion of B causes the motion of A. Children of that age also infer the existence of a common cause of correlated motions when separate interventions on each puppet alone do not stop the motion of the other puppet. That very young children are capable of such inferences suggests that causal learning mechanisms may be fundamental to human cognition.

Dynamic learning

Unlike data-mining programs, human learners forget most of the particular data they receive and must revise their beliefs in light of previous beliefs and new, small samples, and they may have limited processing time. Recent algorithmic work has begun to develop dynamics and low processing requirement for causal learning of Bayes net structures. Spirtes [36] showed that a well-known constraint based learning algorithm [14] can be stopped at any point and yield correct, but less complete, information. Danks [37] has shown that with a single cue, there is a dynamical learning algorithm, analogous to Rescorla–Wagner's, whose equilibria are Cheng's causal powers. Bayesian Bayes nets learners are naturally suited for 'one-step' updating without memory of past data [37].

Alternatives

Kalish and Anh [20] have argued that people make causal judgments based on cognitively available mechanisms connecting putative cause and effect, and have argued that this conception is fundamentally inconsistent with learning from observations, whether passively observed or from interventions. Glymour and Cheng [21] have pointed out that mechanisms are naturally viewed as variables that either intervene between cause and effect, or as other factors that regulate the influence of a cause on an effect, and so have straightforward network representations. They also emphasized the compatibility with inferences based on knowledge of mechanisms with learning from passive observations and from associations resulting from interventions, arguing that unless knowledge of mechanisms is innate it must ultimately be acquired from such data. One must suspect complex cognitive processes in which learned causal relations are generalized and used to provide mechanisms in other, less general, cases. Lien and Cheng's work suggests one such process.

Goldvarg and Johnson-Laird have argued that in adult human understanding causal claims are logical material conditionals, elaborations of the form suggested by Boole and Frege. Traditional objections to this view of content have to do with the monotonicity of logical conditionals and the non-monotonicity of causal claims, and with counterfactuals. 'Striking a match causes a flame' is true, whereas 'Striking a match when there is no oxygen causes a flame' is false, but the material conditional 'If A then B' entails 'If A and C, then B' for any declarative propositions, A, B, and C. Sloman and Lagnado [38] have shown that adult subjects give different responses to 'counterfactual undoing' of premises or conclusions when told that A causes B and, respectively, when told that if A then B, and then asked whether A would obtain in the absence of B, or B would obtain in the absence of A.

Conclusion

The recent work on causal judgment surveyed here opens a raft of experimental and theoretical issues that deserve the attention of cognitive, developmental and mathematical psychologists. Many of the experiments reported here need to be repeated with varied conditions, materials, and methods of subject selection. The individual variation found in several experiments needs to be better explored. The Danks–Mackenzie experiment suggests that a significant subset of subjects can identify causal relations in the absence of a focal set of data in which a potential cause is independent of other causes of the effect, but whether the same phenomena occur in more complex designs, as in Fig. 2, is unknown. Tenenbaum and his collaborators have argued forcefully that in causal learning humans use Bayesian learning procedures, but learning algorithms in causal judgment need further investigation. How particular causal relations, once learned, are generalized and used in subsequent causal inference and in the elaboration of mechanisms – for example, the role of analogy in this process – needs further investigation, continuing the work of Waldmann and of Lien and Cheng. Further algorithmic work suggesting realistic but reliable learning heuristics and procedures is much to be desired. In the longer run, we may hope that this research influences work, and applications, on human factors, where causal judgments have an evident importance. In social psychology, provably correct Bayes net model discovery procedures from statistics and computer science have, as yet, made few inroads against dominant methodologies in which causal models tend to be specified *a priori*, often as regression models or linear structural equation models, and tested by LISREL and similar procedures, without serious exploration of alternative causal models. We may hope that too, changes [39].

References

- 1 Boole, G. (1854) *An Investigation of The Laws of Thought, On Which are Founded the Mathematical Theories of Logic and Probabilities*, Walton and Maberly
- 2 Frege, G. (1984) In *Collected Papers on Mathematics, Logic, and Philosophy* (McGuinness, B., ed.), Blackwell
- 3 Pavlov, I. (1927) In *Conditioned Reflexes; An Investigation of the Physiological Activity of the Cerebral Cortex* (Anrep, G.V., ed.), Oxford University Press/Humphrey Milford
- 4 Piaget, J. (1930) *The Child's Conception of Physical Causality*, Harcourt Brace
- 5 Michotte, A. (1943) *The Perception of Causality*, Basic Books
- 6 Anderson, J. and Sheu, C.F. (1995) Causal inference as perceptual judgements. *Mem. Cogn.* 23, 510–524
- 7 Scholl, B. and Tremoulet, P. (2000) Perceptual causality and animacy. *Trends Cogn. Sci.* 4, 299–309
- 8 Anderson, J. (1983) *The Architecture of Cognition*, Harvard University Press
- 9 Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Theory and Research* (Black, A.H., Prokasy, W.F. eds), pp. 64–99, Appleton-Century Crofts
- 10 Lopez, F. et al. (1998) The rational analysis of human causal and probability judgement. *Rational Models of Cognition* (Oaksford, M., Chater, N., et al. eds), Oxford University Press
- 11 Thelen, E. and Smith, L.B. (1994) *A Dynamic Systems Approach to the Development of Cognition and Action*, MIT Press
- 12 Kiiveri, H. and Speed, T. (1982) Structural analysis of multivariate

- data: a review. In *Sociological Methodology* (Leinhardt, S., ed.), Jossey-Boss
- 13 Pearl, J. (1988) *Probabilistic Reasoning in Uncertain Systems*, Morgan Kaufmann
 - 14 Spirtes, P. et al. (1993) *Causation, prediction and search Springer Lecture Notes in Statistics*, 2nd edn, MIT Press
 - 15 Pearl, J. (2000) *Causality*, Oxford University Press
 - 16 Glymour, C. and Cooper, G. (1999) *Computation, Causation and Discovery*, MIT/AAAI Press
 - 17 Ramsey, J. et al. (2002) Automated remote sensing with near infrared reflectance spectra: carbonate recognition. *Data Mining and Knowledge Discovery* 6, 277–293
 - 18 Shipley, B. (2000) *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference*, Cambridge University Press
 - 19 Heckerman, D.E. et al. (1999) A Bayesian approach to causal discovery. *Computation, Causation and Discovery* (Glymour, C., Cooper, G., et al. eds), pp. 141–166, MIT/AAAI Press
 - 20 Ahn, W. and Kalish, C. (2000) The role of mechanism beliefs in causal reasoning. *Explanation and Cognition* (Keil, F., Wilson, R. eds), MIT Press
 - 21 Glymour, C. and Cheng, P. (1998) Causal mechanism and probability: a normative approach. *Rational Models of Cognition* (Oaksford, M., Chater, N. eds), Oxford University Press
 - 22 Rehder, B. (2001) *A causal model theory of categorization Proceedings of the 21st Annual Conference of the Cognitive Science Society* (Hahn, M., Stoness, S.C. eds), pp. 595–600, Vancouver
 - 23 Gopnik, A. and Nazzi, T. (2001) Words, kinds and causal powers: a theory theory perspective on early naming and categorization. *Early Categorization* (Rakison, D., Oakes, L. eds), Oxford University Press
 - 24 Baker, A. et al. (1993) Selective associations and causality judgements: presence of a strong causal factor may reduce judgements of a weaker one. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 414–432
 - 25 Allan, L. (1980) A note on measurement of contingency between two binary variables in judgement tasks. *Bull. Psychon. Soc.* 15, 147–149
 - 26 Glymour, C. (2001) *The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology*, MIT Press
 - 27 Cheng, P. (1997) From covariation to causation: a causal power theory. *Psychol. Rev.* 104, 367–405
 - 28 Cheng, P. and Novick, L. (2002) *Assessing interactive causal influence* in press
 - 29 Tenenbaum, J.B. and Griffiths, T.L. (2001) Structure learning in human causal induction. In *Advances in Neural Information Processing Systems* (Vol. 13) (Leen, T. et al., eds), MIT Press
 - 30 Waldmann, M.R. and Martignon, L. (1998) A Bayesian network model of causal learning. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (Gernsbacher, M.A., Derry, S.J. eds), pp. 1102–1107, Erlbaum
 - 31 Lien, Y. and Cheng, P. (2000) Distinguishing Genuine from Spurious Causes: a Coherence Hypothesis. *Cogn. Psychol.*
 - 32 Danks, D. (2002) *Equilibria of the Rescorla–Wagner model* in press
 - 33 Lagnado, D. and Sloman, S. (2002) *Learning causal structure Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, George Mason University
 - 34 Hickling, A. and Wellman, H. (2001) The emergence of children's causal explanations and theories: evidence from everyday conversation. *Dev. Psychol.* 37, 668–683
 - 35 Gopnik, A. and Glymour, C. (2002) Causal maps and Bayes nets: a cognitive and computational account of theory-formation. In *The Cognitive Basis of Science* (Carruthers, P. et al., eds), Cambridge University Press
 - 36 Schulz, L. and Gopnik, A. (2001) *Inferring causal relations from observations and interventions Proceedings of the Neural Information Processing Systems Meeting*, Whistler, B.C.
 - 37 Spirtes, P. (2001) *An anytime algorithm for causal inference Proceedings of the Conference on Artificial Intelligence and Statistics*, Fort Lauderdale
 - 38 Sloman, S.A. and Lagnado, D. (2002) *Counterfactual undoing in deterministic causal reasoning Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society*, George Mason University
 - 39 Glymour, C. et al. (1987) *Discovering Causal Structure*, Academic Press

Editor's choice
bmn.com/neuroscience

As a busy cognitive scientist, searching through the wealth of information on BioMedNet can be a bit daunting – the new gateway to neuroscience on BioMedNet is designed to help.

The new gateway is updated weekly and features relevant articles selected by the editorial teams from *Trends in Neuroscience*, *Trends in Cognitive Sciences*, *Current Opinion in Neurobiology* and *Neuron*.

The regular updates include:

- **News** – our dedicated team of reporters from BioMedNet News provides all the news to keep you up-to-date on what's happening – right now.
- **Features** – a special highlight from BioMedNet of one of the latest findings to be reported in the primary journals or at conferences.
- **Conference reporter** – daily updates on the most exciting developments revealed at the Annual meeting for the Society for Neuroscience and other conferences provides a quick but comprehensive report of what you missed by staying home.
- **Mini-reviews and Reviews** – a selection of the best review and opinion articles from the Trends, Current Opinion, and other selected journals.

Why not bookmark the gateway at bmn.com/neuroscience for access to all the news, reviews and informed opinion on the latest scientific advances in neuroscience.