



Bayesian models of cognition

Nick Chater,^{1*} Mike Oaksford,² Ulrike Hahn³ and Evan Heit⁴

There has been a recent explosion in research applying Bayesian models to cognitive phenomena. This development has resulted from the realization that across a wide variety of tasks the fundamental problem the cognitive system confronts is coping with uncertainty. From visual scene recognition to on-line language comprehension, from categorizing stimuli to determining to what degree an argument is convincing, people must deal with the incompleteness of the information they possess to perform these tasks, many of which have important survival-related consequences. This paper provides a review of Bayesian models of cognition, dividing them up by the different aspects of cognition to which they have been applied. The paper begins with a brief review of Bayesian inference. This falls short of a full technical introduction but the reader is referred to the relevant literature for further details. There follows reviews of Bayesian models in Perception, Categorization, Learning and Causality, Language Processing, Inductive Reasoning, Deductive Reasoning, and Argumentation. In all these areas, it is argued that sophisticated Bayesian models are enhancing our understanding of the underlying cognitive computations involved. It is concluded that a major challenge is to extend the evidential basis for these models, especially to accounts of higher level cognition. © 2010 John Wiley & Sons, Ltd. *WIREs Cogn Sci* 2010 1 811–823

INTRODUCTION

From the point of view of the brain, nothing is certain. Sensory input is noisy and extremely partial: the structure of the environment must tentatively be inferred from unreliable scraps of information. Memory is also subject to distortion and interference; and our view of the past thus requires inferring a rich structure on the basis of a sketchy and unreliable record. Linguistic input is notoriously ambiguous, underspecified, may be deliberately deceptive, and its significance can only be a matter of conjecture, rather than certainty. This uncertainty concerning what to believe is paralleled in similar, and equally severe, uncertainties concerning what we want, and how we should act. Yet the brain copes with such uncertainties with surprising

ease—the external world, our memories of the past, and the meaning of people’s utterances, seem, introspectively at least, to be hearteningly stable. It is only in the light of careful experimental analysis that the frailty of such knowledge is revealed—so that perceptual illusions,¹ the unreliability of judgement and memory,² and the slipperiness of linguistic interpretation³ seem, from an introspective point of view, rather unexpected. From this perspective, a fundamental information processing task of the brain is to weld scraps of information together to produce an integrated model of the external world; and to use this model to help determine action and choice.

How can this be done? The Bayesian approach to cognition seeks to model this information processing problem using the mathematical calculus of uncertain inference: probability theory. Each conjecture about the world is associated with a numerical degree of belief, defined to be on the interval between 0 and 1, where 1 corresponds to absolute certainty that the belief is true; and 0 corresponds to absolute certainty that it is false. These beliefs can be identified with probabilities; and a consistent cognitive agent is required to obey the rules of probability theory—at least, if the agent is to avoid paradoxical conclusions.

This probabilistic perspective on the mind can be traced back to one of the origins of probability theory.

*Correspondence to: n.chater@ucl.ac.uk

¹Department of Cognitive, Perceptual and Brain Sciences and Centre for Economic Learning and Social Evolution (ELSE), UCL London, UK

²Department of Psychological Science, Birkbeck College, University of London, London, UK

³School of Psychology, Cardiff University, Cardiff, UK

⁴Psychology Department, University of California, Merced, CA, USA

DOI: 10.1002/wcs.79

Indeed, the very title of Bernoulli's⁴ seminal book *Ars Conjectandi*, 'The Art of Conjecture', embodies the idea that probability captures how people actually make conjectures; as well as providing a calculus for helping people to make conjectures more accurately. Thus, one important strand in the development of probability theory viewed it directly as a theory of thought, as well as a helpful mathematical calculus.

The probabilistic approach can be adopted at three different levels, corresponding to Marr's⁵ three levels of explanation. *Computational* level explanation aims to specify the nature of the problem that the brain faces: the goals of the system and the structure of the environment in which these goals must be achieved. At the computational level, then, probabilistic methods are used to specify the problem that the brain faces. Thus, learning to control an arm, or use a language, might be viewed as problems of probabilistic inference, given certain prior assumptions; and in the light of data gleaned from experience. Modern engineering, machine learning, and artificial intelligence typically view a wide range of information processing problems faced by the brain, from motor control, to speech perception, to object recognition from this probabilistic perspective.

Algorithmic level explanation requires specifying the representations and computational operations over those representations that constitute cognition. Even if the brain faces probabilistic challenges, it may be that it solves them, using some set of heuristics or approximations which do not involve actually carrying out probabilistic calculations. On the other hand, though, the modern technology of probabilistic inference, as explored in state-of-the-art engineering and artificial intelligence systems, does provide a rich set of hypotheses about human cognition. Cognitive science is, after all, a process of reverse engineering; and reverse engineering inevitably draws on the best engineering solutions to the information processing problems that the brain faces.

Finally, even if the brain is probabilistic at the computational and algorithmic levels, this does not necessarily imply that it is probabilistic at the third of Marr's levels of explanation, the *implementational* level. Indeed, probabilistic algorithms used in speech engineering or computer vision run on the binary logic of digital computers. But some neuroscientists have begun to conjecture that the brain may be probabilistic at its very foundations—that individual neurons may convey probabilistic information, that neural populations may capture probability distributions, that basic neural processes might be understood as directly carrying out elementary probabilistic inference.⁶

After providing a brief overview of Bayesian inference, in this rest of this article we survey some of the burgeoning research applying Bayesian models to cognition and perception. Seven sections cover Bayesian models in Perception, Categorization, Learning and Causality, Language Processing, Inductive Reasoning, Deductive Reasoning, and Argumentation.

BAYESIAN INFERENCE

From a probabilistic standpoint, beliefs are a matter of degree. Each hypothesis, H_i , can be associated with a degree of belief $P(H_i)$; and very modest consistency constraints require that these degrees of belief must obey the laws of probability. Thus, the probability distribution over the various H_i can be viewed as characterizing prior beliefs. Suppose that H_i has implications for the data we expect to encounter (e.g., H_i states that the floodlights are on; which if true, makes sense sensory inputs—roughly, the bright ones—more likely than others). These implications can be captured by $P(D|H_i)$, the probability of the data, given the hypothesis. In the light of D , we need to update the priors $P(H_i)$ to $P(H_i|D)$, the probabilities of the hypotheses, given that the data is known. A simple identity of probability theory, Bayes' theorem, shows how this can be done:

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D)}$$

The probability of the data $P(D)$ is not, of course, known independently of the hypotheses that might generate that data—so in practice $P(D)$ is typically expanded using the probabilistic identity:

$$P(D) = \sum_j P(D|H_j)P(H_j)$$

Because of the centrality of the problem of updating beliefs in the light of new information, Bayes' Theorem has very broad application, so much so, indeed, that the interpretation of degrees of belief in terms of probabilities is often known as the Bayesian approach.

If we quantify 'degrees of belief' numerically, as the Bayesian approach presupposes, why should the laws of probability theory, rather than some other principles, define the calculus of degrees of belief? From the point of view of cognitive science, there are two strong arguments for adopting a probabilistic approach. The first, mentioned above, is that violation of the laws of probability leads to paradoxical conclusions. Indeed, the laws of probability can be derived from a variety of plausible, modest, but

very different assumptions concerning how degrees of belief should behave. Perhaps the best known such derivation is the Dutch book theorem,⁷ which shows that, under fairly general conditions, gamblers whose degrees of belief violate the laws of probability will happily accept a combination of bets which are, nonetheless, guaranteed to lose money, whatever their outcomes—which appears to be an unequivocally irrational choice. This type of argument suggests that, given that brains reason spectacularly well about uncertainty, it is unlikely systematically depart from the norms of good probabilistic reasoning by too much—any good uncertain reasoner is, the argument might go, to some degree a good Bayesian, that is, probabilistic, reasoner.

In addition to this a priori line of argument, and perhaps more persuasive from the point of view of the practicing neuroscientist and cognitive scientist is that the Bayesian approach is widely used in engineering approaches to solving the types of problem faced by the brain. Thus, the fields of computer vision, speech recognition, computational linguistics, robotics, machine learning, information retrieval and expert systems, and many more, have seen a dramatic upsurge in the application of probabilistic methods. To the extent that the project of understanding the mind/brain is reverse engineering, that is, attempting to find the engineering principles that underpin neural and cognitive function, then any credible scientific theory has to be good engineering; and the Bayesian approach seems plausibly to pass this test.

Below, we briefly describe the Bayesian approach to cognition in a number of domains, ranging from perception to learning about causal relations, to Bayesian models of higher-level reasoning and argumentation.

PERCEPTION

From a computational level perspective, the problem of perception is that of inferring the structure of the world from sensory input. This problem may seem to be ill-posed, because any given sensory input may have been generated by an infinity of possible states of the world.⁸ From a probabilistic perspective, the infinity of possible interpretations is not in itself problematic. Rather, the challenge of probabilistic inference in perception is to assign probabilities to each of these possible interpretations, based not only on sensory input itself, but prior knowledge. This is a problem of Bayesian inference par excellence.

The Bayesian approach in perception has its beginnings in Helmholtz's⁹ notion of 'unconscious inference', although he did not explicitly use Bayes'

rule.¹⁰ More recently, this perspective has become increasingly influential throughout the brain and cognitive sciences, as well as in computer vision. Moreover, the Bayesian approach is consistent with a broader tradition in perceptual research, the idea that perception is analysis-by-synthesis.¹¹ That is, the perceptual data is presumed to be analyzed (i.e., calculating $P(H|D)$) from a knowledge of the perceptual data that would be generated by various possible scene interpretations (i.e., from a knowledge of $P(D|H)$, and of course a prior distribution $P(H)$ over the hypotheses concerning the scenes)—a transformation which requires the application of Bayes' theorem. In practice, the process of finding an interpretation from which the perceptual data can reasonably be generated requires a combination of bottom-up and top-down perceptual inferences,¹² a process that can be captured computationally by recent methods such as Data-Driven Markov Chain Monte Carlo.¹³ Thus, the Bayesian approach to perception requires that the perceptual system is able to generate sensory input, as well as being able to perceive it; and hence provides a natural explanation of the existence of imagery, consistent with some existing psychological theories,¹⁴ and with experimental data indicating the influence of top-down perceptual processes.¹⁵

Bayesian models of perception have been subjected to direct experimental test in a number of domains (e.g., the integration of sensory cues¹⁶). And a wide variety of computational models of empirical findings in perception have been put forward, ranging from low-level image interpretation,¹⁷ shape from shading,^{8,18} and shape from texture,¹⁹ to boundaries interpolation.^{20,21} There has also been an explosive growth in theories in the field of computational neuroscience which view specific neural mechanisms as carrying out probabilistic computations, from lateral inhibition in the retina,²² to the activity of single cells in the blow-fly,²³ or to populations of neurons including the accumulation of sensory evidence.⁶

Indeed, it turns out that a large class of apparently non probabilistic models of perception can also be accommodated into the Bayesian framework. A long tradition in perception, often viewed as standing in direct opposition to the Bayesian approach, is based on simplicity: the perceptual system is assumed to choose an interpretation of sensory input that provides a briefest encoding of the sensory data. Here, the starting point for the perceiver is a coding language: a representational system in which scenes, and the sensory inputs that they deliver, can be represented. According to simplicity-based explanations, for example, Gestalt principles, such as common fate (grouping

objects with the same movement together, such as a flock of birds) or good continuation (assuming alignment between items, even when occluded, typically indicates they should be grouped, or perhaps part of the same object, for example when the outline of an animal is seen through dense foliage), arise because of a preference for simple codes—codes which specify a single motion direction for the entire flock, rather than for each bird individually; or specify the position of a single occluded object, rather than independently coding the positions of each object fragment. Yet it turns out that simplicity-based approaches to perception^{24–31} are mathematically equivalent to the Bayesian approach, under mild conditions.³² The choice of coding language can be viewed as implicitly specifying a prior probability distribution—such that items that have a brief representation in the language have relatively high prior probability.

CATEGORIZATION

Understanding perceptual input involves the creation of categories. Categorization allows generalization from one category member to another; and also allows the formulation of abstract relations defined over categories, rather than concrete items. From a formal point of view, categorization is an aspect of high-level perception, where categorization of the items is in the scene is just one of many pieces of information that must be recovered from sensory input. In cognitive psychology, early theories of categorization focused on supervised categorization—that is, learning a category from a set of examples, labeled with their category. The two main theoretical approaches both focused on similarity between the item to be classified to a prototypical category exemplar,^{33,34} or alternatively to one of a set of category exemplars.³⁵ While initially formulated in probabilistic terms, both types of theory have increasingly been formulated from a Bayesian point of view.^{36–40} Roughly speaking, the prototype view of categorization can be viewed as assuming that categories corresponds to the Gaussian (or similar) blobs, which may potentially overlap, in some feature space; and the problem of categorization is to work out, given an item, the probability distribution over the Gaussian blobs that may have generated it. According to the simplest formulation, we assume that the participant is certain that the new item is generated by one of the previous encountered categories; but in reality, of course, it is possible that a new item is generated by a category that has not been previously encountered. Thus one extension of the prototype approach, from a probabilistic point of view, is to allow that, in response to a

new item, an agent may postulate a new category; and therefore that the number of categories may grow, perhaps unboundedly, as the number of items categorized increases. This type of ‘nonparametric’ categorization model is widely used in Bayesian models of categorization, from Anderson⁴¹ through to Griffiths et al.⁴² and Goodman et al.⁴³. Exemplar models can then be seen as a limiting case of this class of model.⁴⁴

Viewing the problem of categorization as a matter of probabilistic inference provides more than an interesting notational variant of initial non probabilistic formulations. On the one hand, it provides a fresh perspective on the explanation for classic psychological data. So, to take a simple example, the finding that people are usually able to classify more typical category members more rapidly than less typical category members³⁴ has a natural interpretation: that the features of prototypical items provide more unequivocal evidence for the specific category membership than do less prototypical items; and hence fewer such features needs to be processed, on average, for a category judgment to be made reliably. Moreover, the probabilistic framework provides a starting point for a wide range of generalizations, which may take account of the fact, for example, that a single item may be a member of multiple categories⁴⁵; that the prior assumptions that underpin categorization may be powerfully influenced by background theories⁴⁶; or that the relative importance of different features, and even the choice of appropriate features, may itself depend on the category being considered, and have to be learned.⁴⁵

LEARNING AND CAUSALITY

Conditioning in animals has traditionally been conceived as a matter of the formation of associations, which might be presumed to form on the basis of, for example, the constant conjunction of two events, or their spatial and temporal proximity. Nonetheless, a wide variety of empirical findings has indicated that the animal may be viewed as an intelligent problem solver,⁴⁷ attempting to figure out the structure of the world, from available contingency data. Thus, for example, the discovery of blocking,⁴⁸ that once an animal has learned that an outcome is predicted by one cue, it is less liable to associate that outcome when the second cue is added; however reliable that second cue may be, may suggest that the animal already has an ‘explanation’ of the outcome; and hence no further explanation, for example, in terms of the second cue, is required. To the extent that

the animal is regarded as making inferences about the structure of the environment from observed data concerning the arrival of lights, tones, food pellets, or shocks, the problem that the animal faces appears closely analogous to the general problem of scientific inference, and hence to be naturally modeled with a Bayesian framework.^{49–51} From this point of view, well-known conditioning phenomena, such as that a contingency that has been reliably reinforced is extinguished more rapidly than a contingency that has been partially reinforced, has a natural probabilistic explanation. If a contingency is typically reliable, then after a few ‘extinction’ trials, there is already strong evidence that the state of the world has changed and that the strong tendency is no longer in operation; on the other hand, if the contingency is initially unreliable, then a few such trials are to be expected by chance, in a case, and hence the animal will be slower to reach the conclusion that the world has changed, and that the contingency is no longer in operation. This type of phenomenon is difficult to account for according to some mechanistic associative accounts, because the association formed by partial reinforcement is simply assumed to be weaker, and for this reason should be expected to be eliminated more rapidly.

Similarly, a variety of probabilistic models have been put forward to explain human judgment of contingency and causality, when learning from experience. Cheng,⁵² for example, has put forward a ‘probabilistic contrast’ model of human causal judgment, according to which the strength of a causal relationship is assumed to be measured by the contrast between probability of the effect, in the presence of the cause, and the probability of the effect in the absence of the cause. Griffiths and Tenenbaum⁵³ have proposed a Bayesian model in which the existence of, and the nature of, a potential causal relationship between events is itself inferred from the observed data. This account aims to explain empirical data concerning both how the structure of causal relationships can be learned, as well as the strength of those relationships, which is the primary concern of Cheng’s model. Sloman and Lagnado,⁵⁴ moreover, have directly studied the role of *intervention* in human causal judgments.

According to many standard philosophical accounts of causality, the existence of *causal* relation between two events A and B depends on counterfactual claims about whether, for example, B would still have occurred even if A had been ‘blocked’, leaving everything else unchanged as far as possible. Thus, for example, pressing the ‘alarm set’ button on the alarm clock appears to be causally related to the alarm clock

going off many hours later, in view of our belief that, had the button not been pressed, the alarm would not have sounded. On the other hand, we do not assume that alarm clock sounding is caused by the chiming of the church clock next door, even if this regularly occurs very few seconds before, because we know that if some intervention occurred to stop the church clock is chiming, the alarm sounds nonetheless. It turns out that it is possible to construct a calculus of causal intervention within a probabilistic framework^{55,56}; and there has been recent experimental work attempting to determine how far this framework can provide a useful model of human causality judgments, when intervention is allowed.⁵⁴

Finally, there has been a very promising line of research in cognitive development, exploring Bayesian network models of contingency learning, causal learning, and learning from intervention, throughout development.⁵⁷ For example, Gopnik et al.⁵⁷ discuss a variety of experiments,^{58,59} which demonstrate that pre-school children have the ability to learn causal structures. In particular, this knowledge can be revealed by the nature of the interventions children choose to perform on the experimental apparatus embodying the causal relationships. This knowledge is independent of the frequency information available in the experimental set up and does not appear to be learnable within non-Bayesian frameworks.

Note though, that contingency is a relatively weak source of information about causal relationships. In observing the relationship between an object and its shadow, for example, the fact that the shadow has roughly the same shape as the object that casts it, that the shadow moves predictably when the object moves, and that, in many cases at least, the shadow and object connect smoothly at the object base, provide powerful indications of the existence of a relationship between the two; a trail of footprints in the sand can reasonably be causally attributed to the recent passage of feet purely in virtue of their shape and arrangement. Indeed, a variety of classic psychological demonstrations of ‘perceptual’ causality⁶⁰ and even causal relations underpinned by social interactions,⁶¹ appear to be perceived essentially instantaneously, without requiring prior learning. A strength of the Bayesian approach is that it is, in principle, possible to build models which include richer representations of the physical structure of the environment, or prior knowledge about other aspects of the physical and social world, such that examples of this kind can readily be captured. Such work is at an early stage⁶²; but, for example, there has already been significant progress in constructing computational models of the

attribution of intentions to an agent, from observing the agent's behavior.⁶³

LANGUAGE PROCESSING

Probabilistic approaches have also been influential in recent accounts of language processing and acquisition.⁶⁴ Within linguistics, it has been standard to view probabilistic aspects of language as of marginal importance, although mainly the study of syntax. Language is often viewed as a set of well-formed strings, which are generated by a symbolic grammar, and associated, through systems of symbolic rules, with phonological and semantic representations. The mappings between phonology, syntax, and semantics can be fully described, according to this point of view, without reference to probabilities. Probability is, nonetheless, fundamentally involved in language processing and acquisition in a number of ways.

Notice, for example, that the problem of analog-to-digital conversion, that is, turning an extremely rich and complex acoustic waveform into a discrete phonological representation is an enormously challenging problem of uncertain inference. The speech wave is typically highly locally ambiguous, and can only be disambiguated by piecing together large numbers of locally ambiguous cues, together with background knowledge concerning the speaker, the topic being discussed, and so on. Unsurprisingly, speech technology draws on a rich repertoire of probabilistic methods including hidden Markov models, and neural networks.⁶⁵ Probability plays a similar role in helping to construct a globally coherent parse (and associated semantic representation), in the light of the notorious local ambiguity of natural language, whether such ambiguity is lexical (e.g., *bank* as financial institution or geographical feature), syntactic [e.g., *I saw the man (with the telescope)* vs. *I saw (the man with the telescope)*], or semantic (e.g., *all the witnesses saw a burglar running from the scene*, which might or might not be interpreted as implying that each witness all the same burglar). Again, a globally coherent parse and interpretation of a sentence can only be achieved by integrating these locally ambiguous cues, together with relevant background knowledge; and, just as in the problem of perception, the natural framework in which to consider such integration is probabilistic inference.

Traditional theories of parsing have not, however, taken a probabilistic standpoint; indeed, such accounts have often, instead, focus purely on structural features of the competing parses.⁶⁶ Research over the last decade and a half has, however,

increasingly suggested that a probabilistic integration of multiple cues is used by the language processing system in order to determine the most probable parse and interpretation of the input.^{67–69}

As with other aspects of learning, it is also natural to view the problem of acquiring a language as an example of uncertain inference. Any finite set of linguistic data available to the child will be compatible with an infinite number of languages; and the child must learn to generalize from the observed input to be able to successfully produce and understand linguistic material that has never previously been encountered.

From a non probabilistic point of view, the problem of learning a language appears almost insuperably difficult; it will, for example, be extremely hard for the learner to distinguish between, say, normal English and a version of English with one additional constraint, for example, that it is not grammatically acceptable to begin and end a sentence with the word *fish*, to include more than five adjectives in a noun phrase, or to use a sentence whose sequence of words forms a palindrome (disallowing *dogs chase dogs*). These possible variants of English would be extremely difficult to rule out, because the structures that they disallow are extremely rare, and might not be expected to occur more than a few times, if at all, during childhood. From a probabilistic point of view, these variations need not be ruled out unequivocally, but rather assigned a very low prior probability (e.g., on the basis that prior probability should be inversely related to complexity); from a non probabilistic point of view, such possibilities either need to be ruled out entirely, or pose genuine problems for the learner. Note, though, that languages *do* exhibit numerous apparently arbitrary constraints, which learners are able to successfully learn. So, for example, the child must infer that, while it is acceptable to say *I made the clock break*, *I broke the clock*, and *I made the clock disappear*, it is not acceptable to say *I disappeared the rabbit*, even though the meaning of this string of words is entirely clear. Learning the absence of certain linguistic possibilities has often been viewed as posing 'logical' problems for language acquisition, however, much data the child receives.⁷⁰ From a probabilistic standpoint, it is possible to show that learning is possible in principle, given sufficient data.⁷¹ More important, perhaps, Bayesian analysis of language acquisition provides the tools to assess the prior information that the learner must possess, in order to learn these and other regularities, given realistic estimates of the data available to the child.⁷²

There has, moreover, been increasing interest in building statistical computational models, although not always using a strictly Bayesian framework,

which can potentially model the acquisition of a variety of aspects of phonology, syntax and semantics, ranging from the acquisition of morphology, to syntactic categories, and broad semantic classes^{73–76}; and there has been substantial progress in developing computational models that are able to learn phrase structure and dependency relations from corpora of untagged text.⁷⁷ From the point of view of a Bayesian analysis, the problem of language acquisition remains formidable indeed; but significant progress has been made both in developing specific models of learning, and defining methods for determining what is learnable in principle.

INDUCTIVE REASONING

Inductive reasoning involves drawing conclusions that are probably true, given a set of premises. Consequently, a rational Bayesian approach seems uniquely suited to model induction. Inductive reasoning contrasts with deductive reasoning, in which the conclusion must necessarily follow from a set of premises. In contrast, two inductive arguments can each have some degree of inductive strength (Figure 1).

There is now a well-documented set of empirical regularities on inductive reasoning (see Ref 78, for a more extensive review). These demonstrations all use inference patterns like that in figure 1. Rips,⁷⁹ looked at how people project properties of one category of animals to another (Figure 2(a) and (b)). He found that the more *similar* the premise category is to the conclusion category the stronger the inference (Figure 2a). He also found that the more *typical* the premise category [bluejays (typical) vs. geese (atypical)] the stronger the inference (Figure 2b). Using multiple regression analyses, Rips found distinct contributions of premise-conclusion similarity and premise typicality (see Ref 80 for further investigations of similarity and typicality effects).

Using similar materials, Nisbett et al.,⁸¹ found that participants were very sensitive to the perceived

- Cows have sesamoid bones (a)
All mammals have sesamoid bones
- Ferrets have sesamoid bones (b)
All mammals have sesamoid bones

FIGURE 1 | Inductive arguments vary in strength. The conclusion in argument (a) may seem stronger, or more probable given the evidence, than the conclusion in (b).

- Rabbits have sesamoid bones (a)
Dogs (Bears) have sesamoid bones
- Bluejays (Geese) have sesamoid bones (b)
Blue tits have sesamoid bones
- This Barratos islander is obese (c)
All Barratos islanders are obese
- This Shreeble is blue (d)
All Shreebles are blue
- Cows require vitamin K for the liver to function (e)
Horses require vitamin K for the liver to function
All mammals require vitamin K for the liver to function
- Cows require vitamin K for the liver to function (f)
Ferrets require vitamin K for the liver to function
All mammals require vitamin K for the liver to function

FIGURE 2 | Empirical effects. (a) *Similarity*: when premise and conclusion are more similar (rabbits–dogs) inference is stronger than when they are less similar (rabbits–bears). (b) *Typicality*: typical categories (bluejays) lead to stronger inferences than less typical (geese). *Variability*: variable categories (c) lead to stronger inferences than less variable categories (d). *Diversity*: diverse categories (f) lead to stronger inferences than less diverse categories (e).

variability of the conclusion category. After just one case, variable categories (Figure 2(c)), for example, people on an imaginary island (Barratos) with respect to obesity, lead to weaker inferences than non-variable categories, such as imaginary birds (Shreebles) with respect to color (Figure 2(d)). Nisbett et al.⁸¹ also systematically varied the given number of observations. For example, participants were told that 1, 3, or 20 shreebles had been observed. Inferences were stronger with increased sample size (see also Ref 80). Osherson et al.⁸⁰ showed that diversity of cases also affects inductive strength, that is, Figure 2(f) is considered stronger than Figure 2(e). This diversity effect runs in the opposite direction to the typicality effect: Whereas a typical premise category leads to a fairly strong inductive argument (Figure 2(b)), an argument with two typical premise categories (Figure 2(e)) is weaker than an argument with a typical premise and an atypical premise (Figure 2(f)).

A rational Bayesian model⁸² views evaluating an inductive argument as learning for which categories a property is true or false. In Figure 1(a), the goal is to learn which animals have sesamoid bones. For this novel property, hypotheses must be derived from prior knowledge about familiar properties. People know some facts that are true of all mammals (including cows), but they also know some facts that are true of cows but not some other mammals. The question is which of these known kinds of properties does the novel property, ‘has sesamoid bones’, resemble most, an all-mammal property, or a cow-only property? Crucially it is assumed that *novel properties follow the same distribution as known properties*. Because many known properties of cows are also true of other mammals, argument Figure 1(a) seems fairly strong.

As well as typicality, a Bayesian model also addresses the other key results in inductive reasoning. Similarity effects arise because given that rabbits have sesamoid bones, it more likely that dogs do rather than bears, because rabbits and dogs share more known properties than rabbits and bears. Diversity effects are also addressed. Figure 2(e) will access many idiosyncratic properties true just of large farm animals and so a novel property of cows and horses may seem idiosyncratic to farm animals. In contrast, Figure 2(f) could not access familiar idiosyncratic properties true of just these two animals, so prior hypotheses must be derived from known properties that are true of all mammals or all animals. We have focused here on a narrow class of inductive inference problems that have been especially well-studied empirically. But recent Bayesian models have analyzed a wide range of inductive problems, which can be naturally formulated and modeled in probabilistic terms.^{83,84}

DEDUCTIVE REASONING

Work on ostensibly deductive reasoning tasks reveals many apparent errors and biases when performance is compared to classical logical standards.⁸⁵ The recent emergence of rational Bayesian models casts this performance in a better light by comparing performance to a probabilistic standard.^{86,87} Such models have been developed in all the three main areas investigated in the psychology of reasoning, conditional inference,⁸⁸ data selection,⁸⁹ and syllogistic reasoning.⁹⁰ The key idea behind them all is that the conditional probability, $P(q|p)$, provides the meaning of conditional statements, *if p then q* (e.g., *if you turn the key then the car starts*), and so $P(\text{if } p \text{ then } q) = P(q|p)$. This latter identity is called *The Equation*.^{91,92} To illustrate

the application of rational Bayesian models in this area, we concentrate on conditional inference which is currently the most researched topic in the area.

Four inference patterns have mainly been studied: two which are logically valid: *modus ponens* (MP) and *modus tollens* (MT), and two fallacies: *denying the antecedent* (DA) and *affirming the consequent* (AC) (Figure 3). Classical logic predicts endorsement of the valid inferences and rejection of the fallacies. However, all four inferences are endorsed above 50% and in the characteristic order: $MP > MT > AC > DA$ ⁹³ revealing a large discrepancy between performance and logical expectations.

The core intuition behind a rational Bayesian model of conditional inference is that it must account for the *non monotonicity* of everyday informal reasoning with conditionals.^{94,95} Classical logic is *monotonic* (Figure 4(a)) and hence is unable to account the ability of additional information to *defeat* previously derived conclusions (Figure 4(b)). The only recourse is to question the premises, e.g., in Figure 4(b) to suggest that *birds fly* is false. But surely, while *defeasible*, this is a very useful generalization that we would not want to reject as false.

The Bayesian approach is to adopt The Equation and to treat conditional inference as Bayesian conditionalization.^{87,88} That is, people are trying to determine the posterior probability of the conclusion, $P_1(\text{flies}(a))$, given they now know that the categorical premise holds with certainty, $P_1(\text{bird}(a)) = 1$ (Figure 4(a)). By Bayesian conditionalization, $P_1(\text{flies}(a)) = P_0(\text{flies}(a)|\text{bird}(a))$, that is, the posterior probability of the conclusion equals the prior conditional probability of the conclusion given the categorical premise. Note that this approach easily handles non monotonicity, for example, $P_0(\text{flies}(a)|\text{bird}(a)) = 0.9$ and $P_0(\text{flies}(a)|\text{bird}(a), \text{Ostrich}(a)) = 0$ are perfectly probabilistically consistent (Figure 4b).

This approach cannot immediately apply to MT and the fallacies because, for example, DA requires knowledge of $P_0(\neg\text{flies}(a)|\neg\text{bird}(a))$ and there is insufficient information in the premises

$$\begin{array}{ll}
 \text{(MP)} & \frac{p \Rightarrow q, p}{\therefore q} & \text{(MT)} & \frac{p \Rightarrow q, \neg q}{\therefore \neg p} \\
 \text{(DA)} & \frac{p \Rightarrow q, \neg p}{\therefore \neg q} & \text{(AC)} & \frac{p \Rightarrow q, q}{\therefore p}
 \end{array}$$

FIGURE 3 | The valid inferences, *modus ponens* (MP) and *modus tollens* (MT), and the fallacies, *denying the antecedent* (DA) and *affirming the consequent* (AC), investigated in conditional inference. These inference schema are to be read that if the list of premises above the line are true so must be the conclusion below the line.

$$\begin{aligned}
 \text{(a)} \quad & \frac{\text{triangle}(x) \Rightarrow 3 - \text{sides}(x), \text{triangle}(a)}{\therefore 3 - \text{sides}(a)} \rightarrow \frac{\text{triangle}(x) \Rightarrow 3 - \text{sides}(x), \text{triangle}(a), \text{red}(a)}{\therefore 3 - \text{sides}(a)} \\
 \text{(b)} \quad & \frac{\text{bird}(x) \Rightarrow \text{flies}(x), \text{bird}(a)}{\therefore \text{flies}(a)} \rightarrow \frac{\text{bird}(x) \Rightarrow \text{flies}(x), \text{bird}(a), \text{Ostrich}(a)}{\therefore \neg \text{flies}(a)}
 \end{aligned}$$

FIGURE 4 | Monotonic (a) and non-monotonic (b) conditional inference by MP. In (a), the additional information, that the particular triangle *a* is red, cannot override the original conclusion that qua triangle, *a* has three sides. In contrast, in (b), the additional information, that the particular bird *a* is an Ostrich does override the original conclusion that qua bird, *a* can fly.

FIGURE 5 | Bayesian conditionalization. P_0 = prior probability, for example, prior to learning that *a* is a bird; P_1 = posterior probability, for example, after learning that *a* is a bird. By Bayesian conditionalization $P_1(\text{flies}(a)) = P_0(\text{flies}(a)|\text{bird}(a))$. Note that (a) and (b) are perfectly probabilistically compatible, that is, Bayesian conditionalization is non-monotonic.

$$\begin{aligned}
 \text{(a)} \quad & \frac{\text{bird}(x) \Rightarrow \text{flies}(x), \text{bird}(a)}{\therefore \text{flies}(a)} \quad \frac{P_0(\text{flies}(x) | \text{bird}(x)) = 0.9, P_1(\text{bird}(a)) = 1}{\therefore P_1(\text{flies}(a)) = 0.9} \\
 \text{(b)} \quad & \frac{P_0(\text{flies}(x) | \text{bird}(x)) = 0.9, P_1(\text{bird}(a)) = 1, P_1(\text{Ostrich}(a)) = 1}{\therefore P_1(\text{flies}(a)) = 0}
 \end{aligned}$$

to determine this probability. This is actually also true of $P_0(\text{flies}(a)|\text{bird}(a))$ for MP, which on the subjective view of probability (see Introductory text) must be determined by reference to global world knowledge via the *Ramsey Test*, that is, add the antecedent, $\text{bird}(a)$, to one’s stock of beliefs, make minimal adjustments to incorporate it, and then read off the probability of the consequent, $\text{flies}(a)$, this is $P_0(\text{flies}(a)|\text{bird}(a))$ (Figure 5). To determine the conditional probabilities for DA, AC, and MT requires the assumption that the priors $P_0(\text{flies}(x))$ and $P_0(\text{bird}(x))$ are also available from global world knowledge. Figure 6 shows how well the Bayesian conditionalization model accounts for the principle data on conditional inference.

to persuade yourself or others of a particular, perhaps controversial, position.⁹⁷ The rational Bayesian approach has been extended to at least some aspects of argumentation.⁹⁸ On this view concern centers on how the premises, *P*, of an argument affect the probability of the conclusion, *C*. If $P(C|P)$ is high then the argument has high *inductive strength*.

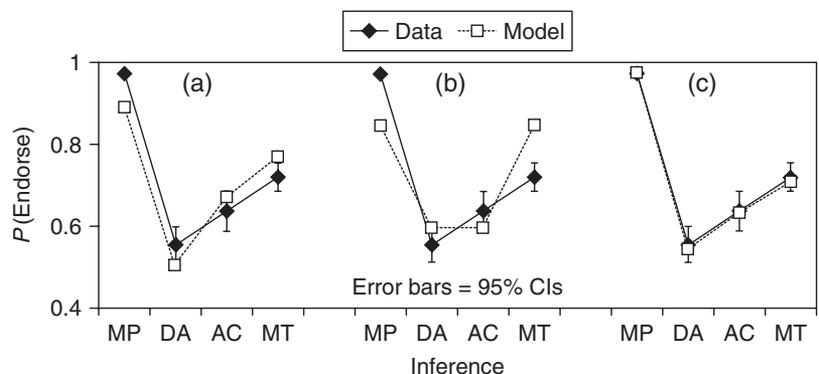
This account has been applied most directly to reasoning *fallacies* in the attempt to understand how some instances seem to be good arguments while others do not.⁹⁹ For example, the classical so-called argument from ignorance, or *argumentum ad ignorantiam*, has many seemingly very weak exemplars:

ARGUMENTATION

Reasoning and decision making often takes place in the service of argumentation, that is, the attempt

Ghosts exist, because nobody has proven that they don’t (1)

FIGURE 6 | Fit of the Bayesian conditionalization model to the empirical data. (a) the fit of the standard account presented in the text; (b) the fit provided by classical logic (modified to incorporate error); (c) the fit of a modified Bayesian conditionalization model.⁹⁶



However, other exemplars of this argument form seem quite strong in scientific and everyday discourse:

This drug is safe, because no-one has found any toxic effects (2)

The classic tool brought to the analysis of fallacies, formal logic, is widely acknowledged to be completely unable to explain the difference between (1) and (2).⁹⁹ More recent *pragma-dialectical* approaches,⁹⁷ which argue that fallacies arise due to the application of rules of argumentation outside the discourse context in which they apply, similarly cannot distinguish (1) from (2). This is simply because (1) and (2) could appear in exactly the same discourse context but (2) would still be regarded as stronger than (1).⁹⁸

The rational Bayesian approach distinguishes (1) and (2) in terms of their inductive strength. Essentially, adequate tests of toxicity exist to establish with a high probability that a drug is safe. However, there are no adequate tests of the non existence or existence of Ghosts that could establish with high probability that Ghosts exist. The Bayesian approach assumes that $P(C|P)$ is calculated by Bayes' theorem which dictates the factors which should influence people's assessments of argument strength.

According to this approach, the argument in (2) corresponds to *negative test validity*, $P(\neg T|\neg e)$, that is, the probability that a drug is not toxic (safe) given there is no evidence of toxicity. This *negative* argument contrast with *positive test validity*, $P(T|e)$ (Figure 7). By Bayes's theorem, these quantities depend on the *sensitivity* and *selectivity* of the test and the *prior belief* that the drug is toxic (Figure 7). If selectivity is higher than sensitivity—a frequent occurrence in real world clinical and psychometric tests—then positive arguments based on $P(T|e)$ are stronger than negative arguments based on $P(\neg T|\neg e)$.^{96,98}

$$P(T|e) = \frac{nh}{nh + (1-l)(1-h)} \quad P(\neg T|\neg e) = \frac{l(1-h)}{l(1-h) + (1-n)h}$$

FIGURE 7 | Positive ($P(T|e)$) and negative ($P(\neg T|\neg e)$) test validity. These probabilities can be calculated from the sensitivity ($P(e|T)$) and the selectivity ($P(\neg e|\neg T)$) of the test and the prior belief that T is true ($P(T)$) using Bayes' theorem. Let n denote sensitivity, that is, $n = P(e|T)$, l denote selectivity, that is, $l = P(\neg e|\neg T)$, and h denote the prior probability of drug A being toxic, that is, $h = P(T)$.

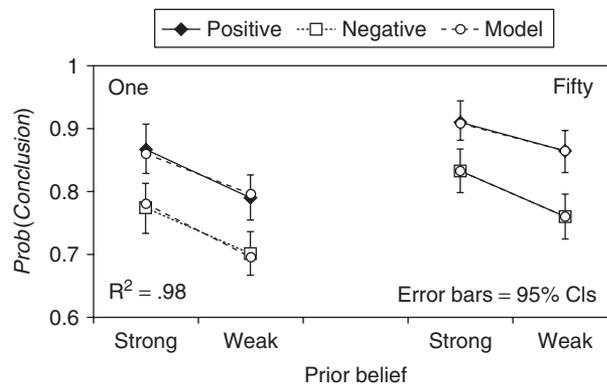


FIGURE 8 | The mean acceptance ratings for Ref 100 by evidence (1 vs. 50 experiments), prior belief (strong vs. weak), and argument type (positive vs. negative). CI = confidence interval, ($N = 84$).

Figure 8 shows the effect of manipulating these factors on peoples' assessments of arguments strength, using an amount of evidence manipulation which should affect sensitivity and selectivity. Fitting the Bayesian model to the data revealed this to be the case: sensitivity and specificity were higher in the high than in the low amount of evidence condition, with $P(\neg P|\neg C) = 0.83$ and $P(P|C) = 0.66$ (high), and $P(\neg P|\neg C) = .77$ and $P(P|C) = 0.46$ (low), respectively. Moreover, sensitivity was higher than selectivity. Similar Bayesian models have also been used to analyze circular reasoning and the slippery slope argument.⁹⁸

CONCLUSION

The brain faces pervasive uncertainty. Bayesian models of cognition aim to understand a wide range of cognitive problems involving uncertainty, ranging from perception to high-level reasoning and argument. Bayesian methods thus may provide a potential link between high-level and low-level cognition that may bridge across each of Marr's levels of explanation. Currently, it would be true to say that the degree of acceptance enjoyed by Bayesian models is roughly inversely related to the level of the cognitive phenomena being modeled, that is, acceptance is greatest at the low neural/perceptual level and decreases as one moves toward higher level phenomena such as reasoning. This seems due in part to availability at the lower level of some quite exquisitely detailed experimental evidence relating the phenomenon to the models. Over the coming years it will be important to see whether similarly detailed and convincing evidence will emerge for Bayesian models of higher level cognition.

REFERENCES

1. Gregory R. *The Intelligent Eye*. London: Weidenfeld and Nicolson; 1970.
2. Johansson P, Hall L, Sikström S, Olsson A. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* 2005, 310:116–119.
3. Quine WVO. *Word and Object*. Cambridge, MA: Harvard University Press; 1960.
4. Bernoulli J. *Ars Conjectandi, The Art of Conjecture*, 1713, (translation and notes by Edith Dudley Sylla, Baltimore, Maryland: John Hopkins University Press, 2005).
5. Marr D. *Vision*. San Francisco, CA: W. H. Freeman; 1982.
6. Beck J, Ma WJ, Kiani R, Hanks T, Churchland AK et al. Probabilistic population codes for Bayesian decision making. *Neuron* 60:1142–52.
7. Lehman R. On confirmation and rational betting. *J Symbolic Logic* 1955, 20:251–262.
8. Freeman WT. The generic viewpoint assumption in a framework for visual perception. *Nature* 1994, 368:542–545.
9. Helmholtz H. *Treatise on Physiological Optics*, vol. 3. New York: Dover; 1910/1962, (English translation by JPC Southall for the Optical Society of America (1925) from the 3rd German edition of *Handbuch der physiologischen Optik* (Hamburg: Voss, 1910; first published in 1867, Leipzig: Voss)).
10. Westheimer G. Was Helmholtz a Bayesian? *Perception* 2008, 37:642–650.
11. Liberman AM, Mattingly IG. The motor theory of speech perception revised. *Cognition* 1985, 21:1–36.
12. Yuille A, Kersten D. Vision as Bayesian inference: analysis by synthesis? *Trends Cogn Sci* 2006, 10:301–308.
13. Tu Z, Zhu S-C. Image segmentation y data-driven Markov chain Monte Carlo. *IEEE Trans Pattern Anal Mach Intell* 2002, 24:657–673.
14. Shepard RN. Ecological constraints on internal representation. *Psychol Rev* 1984, 91:417–447.
15. Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM. Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA* 2006, 103:449–454.
16. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 2002, 415:429–433.
17. Weiss Y. Interpreting images by propagating Bayesian beliefs. In: Mozer, MC, Jordan MI, Petsche T, eds. *Advances in Neural Information Processing Systems 9*. Cambridge MA: MIT Press; 1997, 908–915.
18. Adelson EH, Pentland AP. The perception of shading and reflectance. In: D Knill, W Richards, eds. *Perception as Bayesian Inference*. Cambridge, MA: Cambridge University Press; 1996, 409–423.
19. Blake A, Bulthoff HH, Sheinberg D. Shape from texture: ideal observers and human psychophysics. In: Knill D, Richards W, eds. *Perception as Bayesian Inference*. Cambridge, MA: Cambridge University Press; 1996, 287–321.
20. Feldman J. Bayesian contour integration. *Percept Psychophys* 2001, 63:1171–1182.
21. Feldman J, Singh M. Information along curves and closed contours. *Psychol Rev* 2005, 112:243–252.
22. Barlow HB. (1959)., Sensory mechanisms, the reduction of redundancy, and intelligence. In *The mechanisation of thought processes*, H.M.S.O., London: pp. 535–539.
23. Snippe HP, Poot L, van Hateren JH. A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. *Vis Neurosci* 2000, 17:449–462.
24. Attneave F. Some informational aspects of visual perception. *Psychol Rev* 1954, 61:183–193.
25. Hochberg JE, McAlister E. A quantitative approach to figural “goodness”. *J Exp Psychol* 1953, 46:361–364.
26. Leeuwenberg ELJ. Quantitative specification of information in sequential patterns. *Psychol Rev* 1969, 216–220.
27. Leeuwenberg E. A perceptual coding language for perceptual and auditory patterns. *Am J Community Psychol* 1971, 84:307–349.
28. Leeuwenberg E, Boselie E. Against the likelihood principle in visual form perception. *Psychol Rev* 1988, 95:485–491.
29. Mach E. *The Analysis of Sensations and the Relation of the Physical to the Psychological*. New York: Dover Publications; 1959, (Original work published 1914).
30. Restle F. Theory of serial pattern learning: structural trees. *Psychol Rev* 1970, 77:481–495.
31. Van der Helm PA, Leeuwenberg PA. Goodness of visual regularities: a non- transformational approach. *Psychol Rev* 1996, 103:429–496.
32. Chater N. Reconciling simplicity and likelihood principles in perceptual organisation. *Psychol Rev* 1996, 103:566–581.
33. Reed SK. Pattern recognition and categorization. *Cognit Psychol* 1972, 3:382–407.
34. Rosch E, Mervis CB. Family resemblances: studies in the internal structure of categories. *Cognit Psychol* 1975, 7:573–605.
35. Medin DL, Schaffer MM. Context theory of classification learning. *Psychol Rev* 1978, 85:207–238.
36. Ashby FG, Gott RE. Decision rules in the perception and categorization of multidimensional stimuli. *J Exp Psychol: Learn Mem Cognit* 1988, 14:33–53.

37. Ashby FG, Townsend JT. Varieties of perceptual independence. *Psychol Rev* 1986, 93:154–79.
38. Fried LS, Holyoak KJ. Induction of category distributions: a framework for classification learning. *J Exp Psychol Learn Mem Cogn* 1984, 10:234–257.
39. Lamberts K. Information-accumulation theory of speeded categorization. *Psychol Rev* 2000, 107:227–260.
40. Nosofsky RM. Attention, similarity, and the identification-categorization relationship. *J Exp Psychol Gen* 1986, 115:39–57.
41. Anderson JR. The adaptive nature of human categorization. *Psychol Rev* 1991, 98:409–429.
42. Griffiths TL, Sanborn AN, Canini KR, Navarro DJ. Categorization as nonparametric Bayesian density estimation. In: Oaksford M, Chater N, eds. *The Probabilistic Mind: Prospects for Rational Models of Cognition*. Oxford: Oxford University Press; 2008.
43. Goodman ND, Tenenbaum JB, Griffiths TL, Feldman J. Compositionality in rational analysis: grammar-based induction for concept learning. In: Oaksford M, Chater N, eds. *The Probabilistic Mind: Prospects for Rational Models of cognition*. Oxford: Oxford University Press. 2008.
44. Rosseel Y. Mixture models of categorization. *J Math Psychol* 2002, 46:178–210.
45. Heller KA, Sanborn A, Chater N. Hierarchical learning of dimensional biases in human categorization. *Neural Inf Process Syst*, 2009.
46. Tenenbaum JB, Griffiths TL, Kemp C. Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn Sci* 2006, 10:309–318.
47. Dickinson A. *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press; 1980.
48. Kamin LJ. “Attention-like” processes in classical conditioning. In: Jones MR, ed. *Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation*. Coral Gables, Florida: University of Miami Press; 1968, 9–31.
49. Courville AC, Daw ND, Touretzky DS. Bayesian theories of conditioning in a changing world. *Trends Cogn Sci* 2006, 10:294–300.
50. Gallistel CR, Gibbon J. Time, rate, and conditioning. *Psychol Rev* 2000, 107:289–344.
51. Kakade S, Dayan P. Acquisition and extinction in autoshaping. *Psychol Rev* 2002, 109:533–544.
52. Cheng PW. From covariation to causation: a causal power theory. *Psychol Rev* 1997, 104:367–405.
53. Griffiths TL, Tenenbaum JB. Structure and strength in causal induction. *Cognit Psychol* 2005, 51:354–384.
54. Sloman SA, Lagnado DA. Do we do? *Cogn Sci* 2005, 29:5–39.
55. Spirtes P, Glymour C, Scheines R. *Causation, Prediction and Search*. Cambridge, MA: MIT Press; 1993.
56. Pearl J. *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press; 2000.
57. Gopnik A, Glymour C, Sobel DM, Schulz LE, Kushnir T et al. A theory of causal learning in children: causal maps and Bayes nets. *Psychol Rev* 2004, 111:3–32.
58. Gopnik A, Sobel DM, Schulz LE, Glymour C. Causal learning mechanisms in very young children: two-, three-, and four-year olds infer causal relations from patterns of variation and co-variation. *Dev Psychol* 2001, 37:620–629.
59. Sobel D, Tenenbaum J, Gopnik A. Children’s causal inferences from indirect evidence: backwards blocking and Bayesian reasoning in preschoolers. *Cogn Sci* 2004, 28:303–333.
60. Michotte A. *The Perception of Causality*. New York: Basic Books; 1963.
61. Heider F, Simmel M. An experimental study of apparent behavior. *Am J Community Psychol* 1944, 57:243–259.
62. Sanborn AN, Mansinghka VK, Griffiths TL. A Bayesian framework for modeling intuitive dynamics. In: Taatgen NA, van Rijn H, eds. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society; 2009, 1145–1150.
63. Baker CL, Tenenbaum JB, Saxe RR. Action as inverse planning. *Cognition* In press.
64. Chater N, Manning C. Probabilistic models of language processing and acquisition. *Trends Cogn Sci* 2006, 10:335–344.
65. Rabiner L, Juang L. *Fundamentals of Speech Recognition*. New York: Prentice Hall; 1993.
66. Frazier L. *On Comprehending Sentences: Syntactic Parsing Strategies*. Ph.D. Dissertation, University of Connecticut, 1979.
67. Chater N, Crocker MJ, Pickering MJ. The rational analysis of inquiry: the case of parsing. In: Oaksford M, Chater N, eds. *Rational Models of Cognition*. Oxford: Oxford University Press; 1998, 441–468.
68. Narayanan S, Jurafsky D. A Bayesian model predicts human parse preference and reading time in sentence processing. In: Dietterich TG, Becker S, Ghahramani Z, eds. *Advances in Neural Information Processing Systems*, vol. 14. Cambridge, MA: MIT Press; 2002, 59–65.
69. McRae K, Spivey-Knowlton MJ, Tanenhaus MK. Modeling the influence of thematic fit (and other constraints) in online sentence comprehension. *J Memory Lang* 1998, 38:283–312.
70. Pinker S. Formal models of language learning. *Cognition* 1979, 7:217–283.
71. Chater N, Vitányi P. ‘Ideal learning’ of natural language: positive results about learning from positive evidence. *J Math Psychol* 2007, 51:135–163.

72. Foraker S, Regier T, Khetarpal N, Perfors A, Tenenbaum J. Indirect evidence and the poverty of the stimulus: the case of anaphoric one. *Cogn Sci* 2009, 33:287–300.
73. Goldsmith J. An algorithm for the unsupervised learning of morphology. *Nat Lang Eng* 2006, 12:353–371.
74. Landauer TK, Dumais ST. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychol Rev* 1997, 104:211–240.
75. Redington M, Chater N, Finch S. Distributional information: a powerful cue for acquiring syntactic categories. *Cogn Sci* 1998, 22:425–469.
76. Griffiths TL, Steyvers M. Finding scientific topics. *Proc Natl Acad Sci U S A* 2004, 101:5228–5235.
77. Klein D, Manning C. A generative constituent-context model for improved grammar induction. *Proceedings of the Annual Conference of the Association of Computational Linguistics (ACL 40)*, University of Pennsylvania, Philadelphia, PA, USA, 2002, 128–135.
78. Heit E. Properties of inductive reasoning. *Psychon Bull Rev* 2000, 7:569–592.
79. Rips LJ. Inductive judgments about natural categories. *J Verbal Learn Verbal Behav* 1975, 14:665–681.
80. Osherson DN, Smith EE, Wilkie O, Lopez A, Shafir E. Category-based induction. *Psychol Rev* 1990, 97:185–200.
81. Nisbett RE, Krantz DH, Jepson C, Kunda Z. The use of statistical heuristics in everyday inductive reasoning. *Psychol Rev* 1983, 90:339–363.
82. Heit E. A Bayesian analysis of some forms of inductive reasoning. In: Oaksford M, Chater N, eds. *Rational Models of Cognition*. Oxford: Oxford University Press; 1998, 248–274.
83. Kemp C, Jern A. A taxonomy of inductive problems. In: Taatgen N, van Rijn H. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates; 2009, 255–260.
84. Kemp C, Tenenbaum JB. Structured statistical models of inductive reasoning. *Psychol Rev* 2009, 116:20–58.
85. Evans J.St.B.T. Heuristics and analytic processes in reasoning. *Br J Health Psychol* 1984, 75:541–568.
86. Oaksford M, Chater N. The probabilistic approach to human reasoning. *Trends Cogn Sci* 2001, 5:349–357.
87. Oaksford M, Chater N. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford: Oxford University Press; 2007.
88. Oaksford M, Chater N, Larkin J. Probabilities and polarity biases in conditional inference. *J Exp Psychol Learn Mem Cogn* 2000, 26:883–889.
89. Oaksford M, Chater N. A rational analysis of the selection task as optimal data selection. *Psychol Rev* 1994, 101:608–631.
90. Chater N, Oaksford M. The probability heuristics model of syllogistic reasoning. *Cognit Psychol* 1999, 38:191–258.
91. Adams EW. The utility of truth and probability. In: Weingartner P, Schurz G, Dorn G, eds. *The Role of Pragmatics in Contemporary Philosophy*. Vienna: Holder-Pichler-Tempsky; 1998, 176–1994.
92. Edgington D. On conditionals. *Mind* 1995, 104: 235–329.
93. Schroyens W, Schaeken W. A critique of Oaksford, Chater and Larkin's (2000) conditional probability model of conditional reasoning. *J Exp Psychol Learn Mem Cogn* 2003, 29:140–149.
94. Oaksford M, Chater N. Against logicist cognitive science. *Mind Lang* 1991, 6:1–38.
95. Oaksford M, Chater N. *Rationality in an Uncertain World*. Hove, England: Psychology Press; 1998.
96. Oaksford M, Chater N. Probability logic and the *Modus Ponens-Modus Tollens* asymmetry in conditional inference. In: Chater N, Oaksford M, eds. *The Probabilistic Mind: Prospects for Bayesian Cognitive Science*. Oxford: Oxford University Press; 2008, 97–120.
97. van Eemeren FH, Grootendorst R. *Argumentation, Communication, and Fallacies: a Pragma-Dialectical Perspective*. Mahwah, NJ: Lawrence Erlbaum Associates; 1992.
98. Hahn U, Oaksford M. The rationality of informal argumentation: a Bayesian approach to reasoning fallacies. *Psychol Rev* 2007, 114:704–732.
99. Hamblin CL. *Fallacies*. London: Methuen; 1970.
100. Oaksford M, Hahn U. A Bayesian analysis of the argument from ignorance. *Can J Exp Psychol* 2004, 58:75–85.