INDUCTIVE LOGIC AND EMPIRICAL PSYCHOLOGY

Nick Chater, Mike Oaksford, Ulrike Hahn and Evan Heit

INTRODUCTION

An inductive logic is a system for reasoning that derives conclusions which are plausible or credible, but are nonetheless not certain. Thus, inductive logic goes beyond the more familiar systems of deductive logic, in which the truth of the premises requires the truth of the conclusions. Thus, from *All people are mortal*, we may deductively infer that *Person A is mortal*, *Person B is mortal*, and so on. But from *Person A is mortal*, *Person B is mortal*, and so on, we can inductively derive, with inevitable uncertainty, that *All people are mortal*. However many instances of the generalization we encounter, it is always possible that there is some counterexample of which we are not yet aware. But inductive inference extends far beyond this type of induction from enumeration.

It can be argued, indeed, that many, and perhaps even almost all, inferences outside mathematics involves uncertain, inductive inference. In everyday life, people are routinely forced to work with scraps of information, whether derived from incomplete and noisy sensory input, linguistic information of uncertain provenance, or uncertain background theories or assumptions. Thus, the human mind seems to be more a matter of tentative conjecture, rather than water-tight argument.

To get a sense of the ubiquity of inductive inference, notice that a successful deductive argument cannot be overturned by any additional information that might be added to the premises. Thus, if we know that *All quadralaterals have angles summing to 360 degrees*, and we know that a specific square is a quadrilateral, then we can infer with certainty that it has angles summing to 360 degrees. Any additional information that we might learn about the square cannot overturn this conclusion — if we subsequently learn that it is a large, red, metal square, we can still conclude that its angles have the same sum. Of course, on learning new information we may come to doubt the premises — for example, if I learn that the "square" has been etched onto a globe, I may come to doubt that it is really a square, in the conventional Euclidean sense, at all; and I may suspect that its angles sum to more than 360 degrees. But, although new information may cast doubt on the premises, it cannot lead us to doubt that the conclusion follows, *if* the premises are true. This property of deductive logic is known as *monotonicity*: i.e., adding premises can never overturn existing conclusions.

Handbook of the History of Logic. Volume 10: Inductive Logic. Volume editors: Dov M. Gabbay, Stephan Hartmann and John Woods.
General editors: Dov M. Gabbay, Paul Thagard and John Woods.

In reasoning about the everyday world, by contrast, *non*monotonicity is the norm: almost any conclusion can be overturned, if additional information is acquired. Thus, consider the everyday inference from *Its raining* and *I am about to go outside* to *I will get wet*. This inference is uncertain — indefinitely many additional premises (*the rain is about to stop*; *I will take an umbrella*; *there is a covered walkway*) can overturn the conclusion, even if the premises are correct. The nonmonotonicity of everyday inference is problematic for the application of logical methods to modelling thought. Nonmonotonic inferences are not logically valid and hence fall outside the scope of deductive logical methods.

In psychology, it is clear that many cognitive processes are non-monotonic. In perception, revealing more information about an object can often change the way in which it is interpreted (e.g., a random dot pattern is seen in depth only when it begins to move [Wallach and O'Connell, 1953]; the greyness of a surface is radically altered when information about its three dimensional orientation in relation to the light source is revealed [Adelson, 1993]; and so on). Moreover, in the field of learning, non-mononoticity is clearly the norm: our grammars, causal models or hypotheses may readily be overturned as new sentences are heard, novel actions are performed, or fresh observations are made.

Inductive logic may also be required to capture verbally stated inferences that are typically viewed as instances of deduction. For example, consider the argument from *if you put 50p in the coke machine, you will get a coke* and *I've put 50p in the coke machine*, to *I'll get a coke*. This argument appears to be an instance of a canonical monotonic logical inference: modus ponens.

Yet in the context of commonsense reasoning, this argument does not appear to be monotonic at all. There are innumerable possible additional factors that may block this inference (power failure, the machine is empty, the coin or the can become stuck, and so on). Thus, you can put the money in, and no can of coke may emerge. Attempting to maintain a logical analysis of this argument, these cases could be interpreted as indicating that, from a logical point of view, the conditional rule is simply false — precisely because it succumbs to counterexamples [Politzer and Braine, 1991]. This is, though, an excessively rigorous standpoint, from which almost all everyday conditionals will be discarded as false. But how could a plethora of false conditional statements provide a useful basis for thought and action? From a logical point of view, after all, we can only make inferences from *true* premises; a logical argument tells us nothing, if one or more of its premises is false. In any event, the scope of deductive logic is highly restricted; and it is clear that many psychological processes, from perception, to learning, to everyday inference, are inductive in character.

Philosophical concerns to uncover a system for reasoning with uncertainty are typically initially concerned with normative questions, e.g., what conclusions can justifiably, if tentatively, be drawn, from given premises? and how can such patterns of uncertain inference be systematized? But, from the point of view of the descriptive problem of understanding how the mind operates, closely related questions arise. After all, dealing with uncertainty is, we might expect, an everyday challenge for cognitive systems, human or animal. But for the cognitive system to deal with uncertainty reliably presumably requires the application of some kind of *method* — i.e., conforming with, perhaps only approximately, some set of principles. Without some such foundation, the question of *why* the cognitive system copes with uncertainty (well-enough, most of the time) is left answered. Any particular instance of uncertain reasoning may, of course, be explained by postulating that the cognitive system follows some special strategy, rather than general inference principles. But the mind is able to deal with a hugely complex and continually changing informational environment, for which special-purpose strategies cannot credibly pre-exist. Thus, to explain the reliable (if partial) success of the inductive leaps observed in human cognition, we should consider the possibility that thought is based on some set of principles of good inductive reasoning — i.e., perhaps thought can be explained by reference to some form of inductive logic.

It turns out, of course, that relatively mild and uncontroversial assumptions about how inductive support should work lead, apparently inexorably, to the probability calculus (e.g., [Fitelson, 2005]). While inductive logic might contain more principles than elementary probability — e.g., principles concerning how to deal with inferential relations between logically complex sentences), it is fairly uncontroversial than inductive logics should include the conventional laws of probability. Thus, in restricted contexts, at least, we may replace the term 'inductive logic' with the term 'probability theory' — and, with some exceptions (such as empirical research on explicit inductive inference outlined below), psychologists primarily talk about probability rather than inductive logic.

From the point of view of empirical psychology, then, the proposal that the mind might, in some sense, embody an inductive logic is generally construed in a relatively restricted way. Thus, early, and now unpopular, theories of inductive logic, which pursued the hope that inductive logic might depend purely on the form of sentences, without reference to the meanings of their non-logical terms, or the state of the world (e.g., [Hempel, 1945; Carnap, 1950]) have been little considered. Moreover, theories in which degrees of inductive support are interpreted in terms of proportions of possible worlds (independent of whether these worlds can be conceived by an individual reasoner) are rarely considered (although some theories of probabilistic reasoning have proposed models which involve counting different types of "mental models," which might be viewed as a psychological analogue to the notion of possible worlds, e.g., [Johnson-Laird *et al.*, 1999]). By far the most psychologically natural perspective on inductive logic is to view inductive support as a matter of *subjective* probability — i.e., the degree of belief, by a particular individual, in a specific proposition. After all, the key psychological question is the dynamics of belief-revision: how does the addition of new information modifies ones prior states of belief. The subjectivist view of probability is, particularly in the psychological and artificial intelligence community, known as the Bayesian approach — simple because the simple probabilistic identity which is Bayes' theorem (discussed below) arises so centrally in the process of belief revision. The extent to which cognition should be viewed as conforming with, or departing from, the principles of probability, i.e., the extent to which a Bayesian view of the mind is productive or misleading, has been a central research theme in empirical research in psychology (e.g., [Edwards, 1954; Kahneman *et al.*, 1982; Gigerenzer, 2002]).

As we noted, only very mild restrictions on how "degrees of belief" *should* behave lead to the conclusion that such degrees of belief can be mapped to the [0,1] interval, and should obey the laws of probability. For example, the celebrated "Dutch book theorem" shows that, under fairly general conditions, any gambler whose subjective probabilities deviate from the laws of probability, however slightly, can be mercilessly exploited — i.e., the gambler will cheerful accept a combination of bets such that, whatever happens, she is certain to lose money. Moreover, there are many such arguments, starting with different normative assumptions, which converge on the assumption that "degrees of belief" should be governed by probability. Thus, if we want to explain how it is that people (and, indeed, animals) are able to cope so successfully with their highly uncertain world, the norms of probability provide the beginnings of an answer — to the extent that the mind reasons probabilistically, the normative justifications that imply that this is the "right" way to reason about uncertainty go some way to explaining how it is that the cognitive system deals with uncertainty with a reasonable degree of success.

Alongside these *a priori* normative arguments stands a more practical reason to take probabilistic models of the mind seriously, which arises from artificial intelligence, and related fields such as computer vision and computational linguistics. Understanding any aspect of the biological world is, to some degree, a matter of reverse engineering — of inferring engineering principles from data. Reverse engineering is, though, of course strongly constrained, in practice, by the range of options offered by current "engineering" technologies. There has been something of a probabilistic revolution in the last two decades in proposals concerning engineering solutions to the types of problems solved by the cognitive system. Probabilistic approaches have been increasingly ubiquitous, and widely used, particularly in the light of technical developments that make complex probabilistic models both formally and computationally more manageable than previously. From knowledge-bases, to perception, to language and motor control, there has been considerable application of sophisticated probabilistic methods (e.g., [Chater and Oaksford, 2008; Chater *et al.*, 2006; Oaksford and Chater, 1998; Pearl, 1988; 2000]).

So we have two reasons to take Bayesian models of the mind seriously — probability is arguably the "right" way to deal with uncertainty; and it proves practically useful in solving cognitively-relevant engineering problems. But how useful does the approach to cognition prove to be in practice? How far do alternative models provide a better account? In precisely what sense, if any, should the mind be viewed as probabilistic? And does the Bayesian perspective immediately collapse, in the light of the fact that people are known to make numerous, and systematic, errors in probabilistic reasoning problems. In this chapter, we sketch the Bayesian, subjectivist view of inductive probability in relation to psychological processes. We then survey the application of Bayesian inductive logic in four key areas: language, inductive inference, reasoning, decision making, and argument. Finally, we consider challenges for the attempt to connect inductive logic and empirical psychology.

1 THE BAYESIAN APPROACH TO COGNITION

The vision of probability as a model of thought is as old as the study of probability itself. Indeed, from the outset of the development of the mathematics of probability, the notion had a dual aspect: serving both as a *normative* calculus dictating how people *should* reason about chance events, such as shipping losses or rolls of a dice, but at the same time interpreted as a *descriptive* theory of how people reason about uncertainty. The very title of Bernouilli's great work, *The art of conjecture* [Bernouilli, 1713], nicely embodies this ambiguity — suggesting that it is both a manual concerning how this art should be practiced; and an outline of how the art is actually conducted. This dual perspective was, indeed, not confined merely to probability, but also applied equally well to logic, the calculus of certain reasoning. Thus Boole's [1958/1854] *The Laws of Thought*, which deals with both logical and probabilistic reasoning, also embodies the ambiguity implicit in its title — it aims to be both a description of how thought works; but also views the laws of thought as providing norms to which reason should conform.

In retrospect, the identification, or perhaps conflation, of normative and descriptive programmes seems anomalous. Towards the end of the nineteenth century, mathematics began to break away from the morass of psychological intuition; and throughout the twentieth century, increasingly formal and abstract programmes for the foundations of mathematics developed, seeming ever more distant from psychological notions. Thus, in the context of probability, Kolmogorov provided an axiomatization of probability in terms of σ-algebras, which views probability theory as an abstract formal structure, with no particular linkage to psychological notions concerning degree of belief or plausibility. Indeed, the idea that mathematics should be rooted in psychological notions became increasingly unpopular, and the perspective of *psychologism* became philosophically disreputable. At a practical level, too, the mathematics and psychology of probability became ever more distant. The mathematics became increasingly formally sophisticated, with spectacular results; but most of this work explicitly disavowed the idea that probability was about beliefs at all. The most popular perspective on probability took the view that probabilities should be interpreted, instead, as limiting frequencies over repeatable events. Thus, to say that the probability of a coin falling heads is $\frac{1}{2}$ is to say something like: in the limit, if this event is repeated indefinitely, the proportion of times that the coin comes up heads will tend towards $\frac{1}{2}$. This frequentist [von Mises, 1957] interpretation of probability aims to separate probability entirely from the beliefs of any particular person observing the coin — the probability is supposed to be a fact about the coin, not about degrees of belief of an observer of the coin.

The premise underlying the Bayesian approach to psychology is that this divorce was somewhat premature — and that, at minimum, a limited reconciliation should be attempted. In particular, the conjecture is that many aspect of thought can be understood as, at some level of approximation at least, embodying probabilistic calculations.

We mentioned above that, normative considerations aside, one appeal of probabilistic models of cognition is that probability has swept into vogue in fields concerned with engineering solutions to information processing problems analogous to those solved by the brain. And this work has overwhelming taken the subjectivist, rather than the frequentist view of probability. One reason for this is that, in many practical applications, the frequentist interpretation of probability simply does not apply — probabilities can only be viewed as expressing degrees of belief (or, more neutrally, degrees of partial information — after all, we may not want to attribute full-blown belief to a simple computational model, or an elementary cognitive process). Thus, in speech recognition or computational vision, each sensory input is enormously complex and will never be encountered again. Hence, there is no meaningful limiting frequency concerning the probability that this image is a photograph of a dog, or a wolf. It definitively is one or the other (the frequencies are 0 or 1 for each category). Similarly, the frequentist interpretation is not appropriate for interpreting uncertainty concerning scientific hypotheses, because, of course, any scientific hypothesis holds, or it does not; and hence limiting frequencies across many trials make no sense. In cases where the goal is to quantify the uncertainty about a state of the world, the uncertainty resides in the computational system (the human or animal brain, the machine learner) attempting to infer the probability. But once we interpret probability as concerning subjective states of belief or information — i.e., once we adopt the *subjective* interpretation of probability — then it is natural to frame the computational challenge of recognizing a word, an animal, or an action, or a scientific hypothesis, purely as a matter of probabilistic calculation. Indeed, according to results such as the Dutch book theorem, mentioned above, once we start to assign degrees of uncertainty to states of any kind, it is mandatory that we use the laws of probability to manipulate these uncertainties, on pain of demonstrable irrationality (e.g., being willing to accept combinations of gambles leading to a certain loss).

In perception, as well as in many aspects of learning and reasoning, the primary goal is working out the probability of various possible hypotheses about the state of the world, given a set of data. This is typically done indirectly, by viewing the various hypotheses about the world as implying probabilities concerning the possible sensory data — i.e., we view these various states of the world as implicitly making claims about the probability of different patterns of data. An elementary identity of probability allows us to relate the probabilities that we are interested in $\Pr(H_i|D)$, the probability that hypothesis H_i is true, given the observed data, D, in terms of the probabilities that are presumed to be implicit in the hypotheses themselves — the probabilities $\Pr(D|H_i)$ of the data, given each H_i. The elementary identity follows immediately from the definition of conditional probability:

$$\Pr(H_i|D)\Pr(D) = \Pr(H_i, D) = \Pr(D|H_i)\Pr(H_i)$$

so that we obtain:

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

which is Bayes' theorem. The probability of the data is not, of course, known independently of the hypotheses that might generate that data — so in practice $\Pr(D)$ is typically expanded using the probabilistic identity:

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

Thus, taking a subjective approach to probability, where states of the world may be viewed as uncertain, from the point of view of an agent, implies that making inferences about the likely state of the world is a matter of probabilistic calculation; and such calculations typically invoke Bayes' theorem, to invert the relationship between hypothesis and data. The prevalence of Bayes theorem in this type of calculation has led to this approach to statistics [Bernado and Smith, 1994], machine learning [Mackay, 2003], and scientific reasoning [Howson and Urbach, 1993] to be known as the Bayesian approach — but the point of controversy is not of course the probabilistic identity that is Bayes' theorem; but rather the adoption of the subjective interpretation of probability. Indeed, in cognitive science, given that almost all applications of probability require a subjective interpretation of uncertainty, the probabilistic approach and the Bayesian approach are largely synonymous.

Levels of probabilistic explanation

Probability is, we have suggested, potentially relevant to understanding the mind/brain. But it can be applied in a range of different ways and at different levels of explanation, ranging from probabilistic analysis of the neural processes in perception and motor control, to normative description of how decision makers should act in economic contexts. But these seem to be explanations at very different levels — and it is worth pausing briefly to consider the range of different levels of analysis at which probabilistic ideas may be applied — and hence to clarify the claims that are (and are not) being reviewed in this chapter.

We suggest that the variety of types of explanation can usefully be understood in terms of Marr's [1982] celebrated distinction between three levels of computational explanation: the *computational* level, which specifies the nature of the cognitive problem being solved, the information involved in solving it, and the logic by which it can be solved (this is closely related to the level of rational analysis, see [Anderson, 1990; 1991a; Anderson and Milson, 1989; Anderson and Schooler, 1991; Oaksford and Chater, 1994; 1998a]); the *algorithmic* level, which specifies the representations and processes by which solutions to the problem are computed; and the *implementational* level, which specifies how these representations and processes are realized in neural terms.

The Bayesian approach has potential relevance at each of these levels. As we have noted, the very fact that much cognitive processing is naturally interpreted as uncertain inference immediately highlights the relevance of probabilistic methods at the computational level. This level of analysis is focused entirely on the nature of the problem being solved — there is no commitment concerning how the cognitive system actually attempts to solve (or approximately to solve) the problem. Thus, a probabilistic viewpoint on the problem of, say, perception or inference, is compatible with the belief that at the algorithmic level, the relevant cognitive processes operate via a set of heuristic tricks (e.g., [Gigerenzer and Todd, 1999; Ramachandran, 1994]), rather than explicit probabilistic computations.

One drawback of the heuristics approach, though, at which we have hinted already, is that it is not easy to explain the remarkable generality and flexibility of human cognition. Such flexibility seems to suggest that cognitive problems involving uncertainty may, in some cases at least, be solved by the application of probabilistic methods. Thus, we may take models such as stochastic grammars for language or vision, or Bayesian networks, as candidate hypotheses about cognitive representation. Yet, when scaled-up to real-world problems, full Bayesian computations are intractable, an issue that is routinely faced in engineering applications. From this perspective, the fields of machine learning, artificial intelligence, statistics, informational theory and control theory can be viewed as rich sources of hypotheses concerning tractable, approximate algorithms that might underlie probabilistic cognition.

Finally, turning to the implementational level, one may ask whether the brain itself should be viewed in probabilistic terms. Intriguingly, many of the sophisticated probabilistic models that have been developed with cognitive processes in mind map naturally onto highly distributed, autonomous, and parallel computational architectures, which seem to capture the qualitative features of neural architecture. Indeed, computational neuroscience [Dayan and Abbott, 2001] has attempted to understand the nervous system as implementing probabilistic calculations; and neurophysiological findings, ranging from spike trains in the blow-fly visual system [Rieke *et al.*, 1997], to cells apparently involved in decision making in monkeys [Gold and Shadlen, 2000] have been interpreted as conveying probabilistic information. Nonetheless, large-scale probabilistic calculations over complex internal representations, and reasonably large sets of data, are typically computationally intractable. Thus, typically, the number of possible states of the world grows exponentially with the number of facts that are considered. Calculations over this exponentially large set of world-states is typically viable only to an approximation. Thus, the mind cannot credibly be viewed as a "Laplacian demon," making complete and accurate probabilistic calculations [Gigerenzer and Goldstein, 1996; Oaksford and Chater, 1998b] — but rather must, at best, be viewed as approximating such calculations, perhaps using some very drastic simplifications. How far it is possible to tell an integrated probabilistic story across levels of explanation, or whether the picture is more complex, remains to be determined by future research.

Why is probability so hard?

The question of levels is important in addressing what may appear to be direct evidence against the application of inductive logic in psychology — research on how people reason explicitly about probability. Describing probabilities as degrees of *belief*, as in the subjectivist interpretation of probability, invites comparison with the folk psychological notion of belief, in which our everyday accounts of each other's behaviour are formed (e.g., [Fodor, 1987]). This in turn suggests that people might reasonably be expected to introspect about the probabilities associated with their beliefs. In practice, people often appear poor at making such numerical judgments; and poor, too, at numerical probabilistic reasoning problems, where they appear to fall victim to a range of probabilistic fallacies (e.g., [Kahneman *et al.*, 1982]). The fact that people can appear to be such poor probabilists may seem to conflict with the thesis that many aspects of cognition can fruitfully be modelled in probabilistic terms.

Yet this conflict is only apparent. People struggle not just with probability, but with all branches of mathematics. Yet the fact that, e.g., Fourier analysis, is hard to understand does not imply that it, and its generalizations, are not fundamental to audition and vision. The ability to introspect about the operations of the cognitive system are the exception rather than the rule — hence, probabilistic models of cognition do not imply the cognitive naturalness of learning and applying probability theory.

Indeed, probabilistic models may be most applicable to cognitive process that are particularly well-optimized, and which solve the probabilistic problem of interest especially effectively. Thus, vision or motor control may be especially tractable to a probabilistic approach; and our explicit attempts to reason about chance might often, ironically, be poorly modelled by probability theory. Nonetheless, some conscious judgments have proven amenable to probabilistic analyses, such as assessments of covariation or causal efficacy [Cheng, 1997; Griffiths and Tenenbaum, 2005; Waldmann, 2008], uncertain reasoning over causal models [Sloman and Lagnado, 2004], or predicting the prevalence of everyday events [Griffiths and Tenenbaum, 2006]. But unlike textbook probability problems, these are exactly the sorts of critical real-world judgments for which human cognition should be expected to be optimized.

The probabilistic turn in the cognitive and brain sciences

We have suggested that probabilistic analysis may be especially appropriate for highly optimized aspects of cognition — i.e., the domains for which it is credible that the brain has some dedicated computational "module" or system of modules (e.g., [Fodor, 1983; Shallice, 1988]). Thus, the probabilistic approach has been widely applied in the areas of perception, motor control, and language, where the performance of dedicated computational modules vastly exceeds the abilities of any artificial computational methods by an enormous margin. Before turning to the main topics of this chapter, the somewhat ill-defined area of "central" cognition, we briefly review the much larger and more extensively developed literatures that apply probabilistic methods to these "modular" domains.

Consider, for example, the problem of inferring the structure of the world, from visual input. There are, notoriously, infinitely many states of the environment that can give rise to any perceptual input (e.g., [Freeman, 1994]) — this is just an example of the standard observation, in the philosophy of science, that theory is underdetermined by data [Laudan and Leplin, 1991]; or in statistics, that an infinite number of curves can fit any particular set of data points (e.g., [Mackay, 1992]). A natural objective of the perceptual system, faced with an infinite number of possible interpretations of a stimulus, is to aim to choose the interpretations which are most likely. From this perspective, perception is a problem of probabilistic inference almost by definition.

The idea that the perceptual system seeks the most likely interpretation can be traced to Helmholtz [1910/1962]. More recently, it has been embodied in the Bayesian approach to visual perception that has become prominent in psychology and in neuroscience. This viewpoint has been backed by direct experimental evidence (e.g., [Gregory, 1970; Rock, 1983]) for the inferential character of perceptual interpretation; and also by the construction of detailed theories of particular aspects of perceptual processing, from a Bayesian perspective, including low-level image interpretation [Weiss, 1997], shape from shading [Freeman, 1994, Adelson and Pentland, 1996], shape from texture [Blake *et al.*, 1996], image segmention, object recognition [Tu *et al.*, Zhu, 2005], and interpolation of boundaries [Feldman, 2001; Feldman and Singh, 2005]. Moreover, the function of neural mechanisms involved in visual perception have also been given a probabilistic interpretation — from lateral inhibition in the retina (e.g., [Barlow, 1959]), to the activity of single cells in the blow-fly [Snippe *et al.*, 2000].

The scope of the probabilistic view of perception may, moreover, be somewhat broader than at might first be thought. Although apparently very different from the likelihood view, the simplicity principle in perception, which proposes that the perceptual system chooses the interpretation of the input which provides the simplest encoding of that input (e.g., [Attneave, 1954; Hochberg and McAlister, 1953; Leeuwenberg, 1969; 1971; Leeuwenberg and Boselie, 1988; Mach, 1959/1914; Restle, 1970; Van der Helm and Leewenberg, 1996], though see [Olivers *et al.*, 2004]) turns out to be mathematically equivalent to the likelihood principle [Chater, 1996]. Specifically, under mild mathematical restrictions, for any probabilistic analysis of a perceptual inference (using particular prior probabilistic assumptions) there is a corresponding simplicity-based analysis (using a particular coding language, in which the code-length of an encoding of perceptual data in terms of an interpretation provides the measure of complexity), such that the most likely and the simplest interpretations co-incide. Thus, theories of perception based on simplicity and coding, and theories of neural function based on decorrelation and information compression (e.g., [Barlow, 1959]) can all be viewed as part of the Bayesian probabilistic approach to perception.

The study of perceptuo-motor control provides a second important area of Bayesian analysis. Sensory feedback, typically integrated across different modalities (e.g., visual and haptic information about the positions of, e.g., the hand), contributes to estimating the current state of the motor system. Knowing this current state, and the location, and layout, various aspects of the external environment, is essential for the brain to be able to plan successful motor movements. The precise way in which movements, such as a grasp, are carried out, is likely to have consequences in terms of "utility" for the agent. Thus, successfully grasping a glass of orange may presage a pleasant drink; a less successful grasp may result in unnecessary delay, a slight spillage, a broken glass, or a stained sofa. The motor system needs to choose actions which, given the precision of the information that it has, and the agent's utilities, gives the best expected outcome. The machinery of Bayesian decision theory [Berger, 1985] can be recruited to address this problem.

Bayesian decision theory has been widely applied as a theoretical framework for understanding the control of movement (e.g., [Koerding and Wolpert, 2006]). A wide range of experimental evidence has indicated that movement trajectories are indeed accurately predictable in these terms. In a particularly elegant study, Koerding and Wolpert [2004a] showed that people rely on prior knowledge, rather than evidence from sensory input, depending on the relative precision of each source of information, in a simple repeated motor task. This suggests that the brain learns to model both the distribution of outcomes in prior trials, and the reliability of sensory input — as performance is accurately tuned to the particular distributions of each to which participants are exposed. Similar effects arise not just in movement trajectories, but in force estimation [Koerding and Wolpert, 2004b] and sensory motor timing [Miyazaki *et al.*, 2005].

This work can be generalized to consider the on-line planning of motor movements — i.e., the brain must plan trajectories so that it's on-line estimation of its own state, and ability to dynamically modify that state, lead to the optimal trajectories. The technical extension of Bayesian methods to problems of this type is the subject of the field of on-line feedback control, and there is experimental evidence that people's movements are well-predicted by these methods (e.g., [Knill and Saunders, 2003; Todorov and Jordon, 2002]). Overall, the Bayesian framework has proved to be a remarkably productive framework in which to analyse human motor control.

We now turn to the main topics of this chapter, the somewhat ill-defined area of "central" cognition. However, we begin with language. Despite being characterised as a modular system [Chomsky, 1981; Fodor, 1983], language is really at the borderline between modular input systems and central systems involved in inference, argument and decision making [Fodor, 1983]. The problems that are solved by central systems are invariably posed linguistically and interact strongly with mechanisms of language interpretation.

2 LANGUAGE

The processing and acquisition of language is a central topic in cognitive science. Yet, perhaps surprisingly, the first steps towards a cognitive science of language involved driving out, rather than building on, probability. Whereas structural linguistics focussed on finding regularities in the statistical complexities of language corpora, the Chomskyan revolution focussed on the abstract rules governing linguistic "competence," based on judgements of linguistic acceptability [Chomsky, 1965]. Whereas behaviorists viewed language as a stochastic process determined by principles of reinforcement between stimuli and responses, the new psycholinguistics viewed language processing as governed by internally represented linguistic rules [Fodor *et al.*, 1974]. And interest in statistical and information-theoretic properties of language [Shannon, 1951] was replaced by the mathematical machinery of formal grammar.

In sum, probability has had bad press in the cognitive science of language. The focus on complex linguistic representations (feature matrices, trees, logical representations) and rules defined over them has crowded out probabilistic notions. And the impression that probabilistic ideas are incompatible with the Chomskyan approach to linguistics has been reinforced by debates which appear to pitch probabilistic and related quantitative/connectionist approaches against the symbolic approach to language [Marcus *et al.*, 1999; Pinker, 1999; Seidenberg, 1997; Seidenberg and Elman, 1997].

The recent development of *sophisticated* probabilistic models, casts these issues in a different light. Such models may be *defined over* symbolic rules and representations, rather than being in opposition to them. Thus, grammatical rules may be associated with probabilities of use, capturing what is linguistically likely, not just what is linguistically possible. From this viewpoint, the probabilistic ideas augment symbolic models of language [Klavans and Rednik, 1996; Manning, 2003].

Yet this complementarity does not imply that probabilistic methods merely add to symbolic work, without modification. On the contrary, the "probabilistic turn," broadly characterized, has led to some radical re-thinking in the cognitive science of language, on a number of levels.

In linguistics, there has been renewed interest in phenomena that seem inherently graded and/or stochastic, from phonology to syntax [Bod *et al.*, 2003; Fanselow *et al.*, 2006; Hay, and Baayen, 2005] — this linguistic work is complementary to the focus of Chomskyan linguistics. There have also been revisionist perspectives on the strict symbolic rules thought to underlie language. Although inspired by a type of probabilistic connectionist network, standard optimality theory attempts to define a middle ground of ranked, violable linguistic constraints, used particularly to explain phonological regularities [Smolensky and Legendre, 2006]. It has been extended to employ increasingly rich probabilistic variants. And in morphology, there is debate over whether "rule+exception" regularities (e.g., English past tense, German plural) are better explained by a single stochastic process [Hahn and Nakisa, 2000].

While touching on these issues, this review explores a narrower perspective: that language is represented by a probabilistic model [Manning, 2003]; that language processing involves generating or interpreting using this model; and that language acquisition involves learning such models. (Another interesting line of work that we do not review assumes instead that language processing is based on memory for past instances, and not via the construction of a model of the language [Daelemans and van den Bosch, 2005]). Moreover, for reasons of space, we shall focus mainly on parsing and learning *grammar*, rather than, for example, exploring probabilistic models of how words are recognized [Norris, 2006] or learned [Xu and Tenenbaum, 2007]. We will see that a probabilistic perspective adds to, but also substantially modifies, modelling the symbolic rules, representations and processes underlying language.

From grammar to probabilistic models

To see the contribution of probability, let us begin without it. According to early Chomskyan linguistics, language is internally represented as a grammar: a system of rules that specify all and only allowable sentences. Thus, parsing is viewed as the problem of inferring an underlying linguistic tree, $t \in T$, from the observed strings of words, $s \in S$. Yet natural language is notoriously ambiguous — there are many ways in which local chunks can be parsed, and exponentially many ways in which these parses can be stitched together to produce a global parse. Searching these possibilities is hugely challenging; and there are often many globally possible parses (many t, for a single s). The problem gets dramatically easier if the cognitive system knows that the bracketing [*the* [*old* [*man*]]] is much more likely than [[*the old*] *man*] (though this latter reading is possible, as in *the old man the boats*). This helps locally prune the search space; and helps decide between interpretations for globally ambiguous sentences. In particular, Bayesian methods specify a framework showing how information about the probability of generating different grammatical structures, and their associated word strings, can be used to infer grammatical structure from a string of words. This Bayesian framework is analogous to probabilistic models of vision, inference and learning; what is distinctive is the specific structures (e.g., trees, dependency diagrams) relevant for language.

In computational linguistics, the practical challenge of parsing and interpreting corpora of real language (typically text, sometimes speech) has led to a strong focus on probabilistic methods. However, computational linguistics often parts company from standard linguistic theory, which focuses on much more complex grammatical frameworks, where probabilistic and other computational methods cannot readily be applied. But computational linguistics does, we suggest, provide a valuable source of hypotheses for the cognitive science of language.

Formally, probabilistic parsing involves estimating $\mathrm{Pr}_m(t|s)$, i.e., estimating the likelihood of different trees, t, given a sentence, s, and given a probabilistic model $\mathrm{Pr}_m$of the language:

$$(1) \quad \Pr_m(t|s) = \frac{\Pr_m(t, s)}{\sum_{t'} \Pr_m(t', s)}$$

The probabilistic model can take as many forms as there are linguistic theories (and linguistic structures, t,may equally be trees, attribute-value matrices, dependency diagrams, etc.). For example, suppose that our grammar is a context-free phrase structure grammar. Probabilities are defined for expanding each node in a tree using a given rule. The product of probabilities in a derivation gives the overall probability of that tree.

A particular syntactic ambiguity, much studied in psycholinguistics, concerns prepositional phrase attachment, e.g., *she saw the boy with the telescope*. The parser has to decide: does the prepositional phrase (e.g., *with the telescope*) modify the verb phrase describing the girl's action i.e., she saw-with-a-telescope the boy; or the noun phrase *the boy* — i.e., she saw the-boy-with-a-telescope? This question is a useful starting point for discussing the role of probability in the cognitive science of language.

Principles, probability, and plausibility in parsing

Classical proposals in psycholinguistics assumed that disambiguation occurs using *structural* features of the trees. For example, the principle of minimal attachment would prefer the first reading, because it has one less node [Frazier and Fodor, 1978]. The spirit of this proposal could, though, be recast probabilistically: the probability of a tree is the product of the probabilities at each node; and hence, other things being equal, fewer nodes imply higher probability.

Structural principles in parsing have come under threat from varied parsing preferences within and across languages. But a stochastic grammar may capture different parsing preferences across languages, because the probability of different structures may differ across languages. A structure with fewer nodes, but using highly improbable rules (estimated from a corpus) will be dispreferred. Psycholinguists are increasingly exploring corpus statistics across languages, and parsing preferences do seem to fit the probabilities evident in each language [Desemet *et al.*, 2006; Desmet and Gibson, 2003].

A second problem for structural parsing principles is the influence of lexical information. Thus, the preference for the structurally analogous *the girl saw the boy with a book* appears to reverse — because books are not aids to sight as telescopes are. The pattern flips back with a change of verb: *the girl hit the boy with a book*, because books can be aids to hitting. The probabilistic approach seems useful here — because it seems important to integrate the constraint that seeing-with-telescopes is much more likely than seeing-with-books.

One way to capture these constraints aims to capture statistical (or even rigid) regularities between head words of phrases. For example, "lexicalized" grammars, which carry information about what material co-occurs with specific words, substantially improve computational parsing performance [Charniak, 1997; Collins, 2003].

Plausibility and statistics

Statistical constraints between words are, however, a crude approximation to what sentences are *plausible*. In an off-line judgement task, we use world knowledge, understanding of the social and environmental context, pragmatic principles, and much more, to determine what people might plausibly say or mean. Determining whether a statement is plausible may involve determining how likely it is to be true; but also whether, given the present context, it might plausibly be *said*. The first issue requires a probabilistic model of general knowledge [Oaksford and Chater, 1998; Tenenbaum *et al.*, 2006]. The second issue requires engaging "theory of mind" (inferring the other's mental states), and invoking principles of pragmatics. Models of these processes, probabilistic or otherwise are very preliminary [Jurafsky, 2003].

A fundamental theoretical debate is whether plausibility is used on-line in parsing decisions. Are statistical dependencies between words used as a computationally cheap surrogate for plausibility? Or are both statistics and plausibility deployed on-line, perhaps in separate mechanisms? Eye-tracking paradigms [Tanenhaus *et al.*, 1995; McDonald and Shillcock, 2003] have been used to suggest that both factors are used on-line, though the interpretations are controversial. Recent work indicates that probabilistic grammar models often predict the time course of processing [Jurafsky, 1996; Narayanan and Jurafsky, 2002, Hale, 2003], though parsing preferences also appear to be influenced by additional factors, including the linear distance between the incoming word and the prior words to which it has a dependency relation [Grodner and Gibson, 2005].

Is the most likely parse favoured?

In the probabilistic framework, it is typically assumed that on-line ambiguity resolution favours the most probable parse. Yet Chater, Crocker and Pickering [1998] suggest that, for a serial parser, whose chance of "recovery" is highest if the "mistake" is discovered soon, this is overly simple. In particular, they suggest that because parsing decisions are made *on-line* [Pickering *et al.*, 2000] there should be a bias to choose interpretations which make *specific* predictions that might rapidly be falsified. For example, after *John realized his...* the more probable interpretation is that realized introduces a reduced relative clause (i.e., *John realized (that) his...*). On this interpretation, the rest of the noun phrase after *his* is unconstrained. By contrast, the less probable transitive reading (*John realized his goals/potential/objectives*) places very strong constraints on the subsequent noun phrase. Perhaps, then, the parser should favour the more specific reading, because if wrong, it may rapidly and successfully be corrected. Chater *et al.* [1998] provide a Bayesian analysis of "optimal ambiguity resolution" capturing such cases. The empirical issue of whether the human parser follows this analysis [Pickering *et al.*, 2000], and even the correct probabilistic analysis of sentences of this type [Crocker and Brant, 2000], is not fully resolved.

Beyond parsing

We have here focussed on parsing. But the "probabilistic turn" applies across language processing, from modelling lexical semantics to modelling processing difficulty. Note, though, that integrating these diverse approaches into a unified model of language is extremely challenging; and many of the theoretical issues that have traditionally concerned psycholinguistics are re-framed rather than resolved by a probabilistic approach.

Probabilistic perspectives on language acquisition

Probabilistic language processing presupposes a probabilistic model of the language; and uses that model to infer, for example, how sentences should be parsed, or ambiguous words interpreted. But how is such a model, or indeed simply a non-probabilistic grammar, acquired? Chomsky [1981] frames the problem as follows: the child has a hypothesis-space of candidate grammars; and must choose, on the basis of (primarily linguistic) experience one of these grammars. From a Bayesian standpoint, each candidate grammar is associated with a prior probability; and these probabilities will be modified by experience using Bayesian updating. The learner will presumably choose a language with high, and perhaps the highest, posterior probability.

The poverty of the stimulus?

Chomsky [1965] influentially argued that the learning problem is unsolvable without strong prior constraints on the language, given the 'poverty' (i.e., partiality and errorfulness) of the linguistic stimulus. Indeed, Chomsky [1981] argued that almost all syntactic structure, aside from a finite number of binary parameters, must be innate. Separate mathematical work by Gold [1967] indicated that, under certain assumptions, learners provably cannot converge on a language even "in the limit" as the corpus becomes indefinitely large (see [Pinker, 1979] for discussion). indexideal@"ideal" learnig

A probabilistic standpoint yields more positive learnability results. For example, Horning [1971] proved that phrase structure grammars are learnable (with high probability) to within a statistical tolerance, if sentences are sampled as independent, identically distributed data. Chater and Vitányi [2007] generalize to a language which is generated by any computable process (i.e., sentences can be interdependent, and generated by any computable grammar), and show that prediction, grammaticality, and semantics, are learnable, to a statistical tolerance. These results are "ideal" however — they consider what would be learned, if the learner could find the shortest representation of linguistic data. In practice, the learner will find a short code, not the shortest, and theoretical results are not available for this case. Nonetheless, from a probabilistic standpoint, learning looks more tractable — partly because learning need only succeed with high probability; and to an approximation (speakers may learn slightly different idiolects).

Computational models of language learning

Yet the question of learnability, and the potential need for innate constraints, remains. Machine learning methods have successfully learned small artificial context-free languages (e.g., [Lari and Youg, 1990]), but profound difficulties in extending these results to real language corpora have led computational linguists to focus on learning from parsed trees [Charniak, 1997; Collins, 2003] — presumably not available to the child. Connectionism is no panacea here — indeed, connectionist simulations of language learning typically use small artificial languages [Elman, 1990; Christiansen and Chater, 2001] and, despite having considerable psychological interest, they scale poorly.

By contrast, many simple but important aspects of language structure have successfully been learned from linguistic corpora by distributional methods. For example, good approximations to syntactic categories and semantic classes have been learned by clustering words based on their linear distributional contexts (e.g., the distribution over the word that precedes and follows each token of a type) or broad topical contexts (e.g., [Schütze, 1995; Redington *et al.*, 1998]). One can even simultaneously cluster words exploiting local syntactic and topical similarity [Griffiths *et al.*, 2005].

Recently, though, Klein and Manning [2002; 2004] have made significant progress in solving the problem of learning syntactic constituency from corpora of unparsed sentences. Klein and Manning [2002] extended the success of distributional clustering methods for learning word classes by using the left and right word context of a putative constituent and its content as the basis of similarity calculations. Such a model better realizes ideas from traditional linguistic constituency tests which emphasize (i) the external context of a phrase ("something is a noun phrase if it appears in noun phrase contexts") at least as much as its internal structure, and (ii) proform tests (testing replacing a large constituent with a single word member of the same category). Klein and Manning [2004] extended this work by combining such a distributional phrase clustering model with a dependency-grammar-based model. The dependency model uses data on word co-occurrence to bootstrap word-word dependency probabilities, but the work crucially shows that more is needed than simply a model based on word co-occurrence. One appears to need two types of prior constraint: one making dependencies more likely between nearby words than far away words, and the other making it more likely for a word to have few rather than many dependents. Both of Klein and Manning's models capture a few core features of language structure, while still being simple enough to support learning. The resulting combined model is better than either model individually, suggesting a certain complementarity of knowledge sources. Klein and Manning show that high-quality parses can be learned from surprisingly little text, from a range of languages, with no labeled examples and no language-specific biases. The resulting model provides good results, building binary trees which are correct on over 80% of the constituency decisions in hand-parsed English text.

This work is a promising demonstration of empirical language learning, but most linguistic theories use richer structures than surface phrase structure trees; and a particularly important objective is finding models that map to meaning representations. This remains very much an area of ongoing research, but inter alia there is work on probabilistic parsing with richer formalized grammar models based on learning from parsed data [Johnson and Riezler, 2002; Toutanova *et al.*, 2005] some work on mapping to meaning representations of simple data sets [Zettlemoyer and Collins, 2005], and work on unsupervised learning of a mapping from surface text to semantic role representations [Swier and Stevenson, 2005].

Poverty of the stimulus, again

The status of Chomsky's poverty of the stimulus argument remains unclear, beginning with the question of whether children really do face a poverty of linguistic data (see the debate between Pullum and Scholz [2002] and Legate and Yang [2002]). Perhaps no large and complex grammar can be learned from the child's input; or perhaps certain specific linguistic patterns (e.g., those encoded in an innate universal grammar) are in principle unlearnable. Probabilistic methods provide a potential way of assessing such questions. Oversimplifying somewhat, suppose that a learner wonders whether to include constraint C in her grammar. C happens, perhaps coincidentally, to fit all the data so far encountered. If the learner does not assume C, the probability of different sentences is, say, $\Pr(x)$. Constraint C only applies to probability mass p of these sentences, where $p = \sum_{x:C(x)} \Pr(x)$. Thus, each sentence obeying C is $1/p$ times more probable, if the constraint is true than if it is not (if we simply rescale the probability of all sentences obeying the constraint). Thus, after n sentences, the probability of the corpus, is $1/p^n$ greater, if the constraint is included. Yet, a more complex grammar will typically have a lower prior probability. If the ratio of priors for grammars with/without the constraint is greater than $1/p^n$, then, by Bayes' theorem, the constraint is unlearnable in n items.

Presently, theorists using probabilistic methods diverge widely on the severity of prior "innate" constraints they assume. Some theorists focus on applying probability to learning parameters of Chomskyan Universal Grammar [Gibson and Wexler, 1994; Niyogi, 2006]; others focus on learning relatively simple aspects of language, such as syntactic or semantic categories, or approximate morphological decomposition, with relatively weak prior assumptions [Redington *et al.*, 1998; Brent and Cartwright, 1996; Landauer and Dumais, 1997]. Probabilistic methods should be viewed as a framework for building and evaluating theories of language acquisition, and for concretely formulating questions concerning the poverty of the stimulus, rather than as embodying any particular theoretical viewpoint. This point arises throughout cognition — although probability provides natural models of learning, it is an open question whether initial structure may be critical in facilitating such learning. For example, Culicover (1999) argues that prior structure over Bayesian networks is crucial to support learning.

Language acquisition and language structure

How far do probabilistic perspectives on language structure, and language acquisition, interact? Some theorists argue that language should not best be described as rules and exceptions, but as a system of graded "quasi-regular" mappings. Notable examples of such mappings including the English past-tense, the German plural, and spelling-to-sound correspondences in English; but a closely related viewpoint has been advocated for syntax [Culicover, 1999; Tomasello, 2003] and aspects of semantics [Baayen and Moscoso del Prado, 2005]. Some theorists argue [Pierrehumbert, 2001] that such mappings are better learned using statistical or connectionist methods, which learn according to probabilistic principles. By contrast, traditional rule-and-exception views are typically associated with non-probabilistic hypothesis generation and test. Nonetheless, we see no necessary connection between these debates on the structure of language, and models of acquisition.

Language: Summary

Understanding and producing language involves complex patterns of uncertain inference, from processing noisy and partial speech input to lexical identification, syntactic and semantic analysis, to language interpretation in context. Acquiring language involves uncertain inference from linguistic and other data, to infer language structure. These uncertain inferences are naturally framed using probability theory: the calculus of uncertainty. Historically, probabilistic approaches to language are associated with simple models of language structure (e.g., local dependencies between words); but, across the cognitive sciences technical advances have reduced this type of limitation. Probabilistic methods are also often associated with empiricist views of language acquisition — but the framework is equally compatible with nativism — that there are prior constraints on the class of language models. Indeed, as we have seen, probabilistic analysis may provide one line of attack (alongside the empirical investigation of child language) for assessing the relative contribution of innate constraints and corpus input, in language acquisition. Overall, probabilistic methods provide a rich framework for theorising about language structure, processing, and acquisition, which may prove valuable in developing, and contrasting between, a wide range of theoretical perspectives.

3 INDUCTIVE REASONING

Historically, in empirical psychology, inductive reasoning has typically been studied separately from deductive reasoning, by separate groups of researchers using different theoretical frameworks. In the next few sections after this one on inductive reasoning, we will review recent attempts to apply inductive logic to psychological studies of deductive reasoning. This will raise the possibility that a unified approach across these diverse reasoning tasks might be achievable. Even in inference tasks that might have a deductive solution people might be more concerned with their inductive strength.

In this section, we concentrate on empirical studies of inductive reasoning, and address the question of whether normative inductive logic can explain the factors on which peoples' judgements of inductive strength depend. In moving to reasoning behaviour, we are now more directly in the realm of central processes and of explicit verbal reasoning tasks of the type dealt with in logic, be it deductive or inductive. Inductive reasoning, in its broadest sense, concerns inference from specific premises to general statements or to other non-logically related specific statements. So, for example, we might be given observations that robins have anatomical feature X, and be asked how likely it is that all birds have anatomical feature X. Or, more usually in experimental tasks, people are asked about the likelihood that eagles or sparrows also have that anatomical feature.

Inductive reasoning involves drawing conclusions that are probably true, given a set of premises. Inductive reasoning can thus be contrasted with deductive reasoning, in which the conclusion must necessarily follow from a set of premises. For example, the following two arguments (1 and 2) each have some degree of inductive strength.

(1) Cows have sesamoid bones.
All mammals have sesamoid bones.

(2) Ferrets have sesamoid bones.
All mammals have sesamoid bones.

Whereas all valid deductive arguments are perfectly strong, inductive arguments can differ in their perceived inductive strength. In the examples above, the conclusion in argument (1) may seem stronger, or more probable given the evidence, than the conclusion in (2)

Inductive reasoning is sometimes characterized as drawing inferences from specific statements to more general statements (as in arguments [1] and [2]), in contrast to deductive reasoning which would run from general statements to specifics. Although there is a grain of truth in this characterization, there is actually a broader variety of deductive as well as inductive arguments (Skyrms, 1977). For example, the following deductively valid argument (3) does not draw a more specific inference from general statements:

(3) Gorillas are apes.
Apes are mammals.
Gorillas are mammals.

Likewise it would be possible to draw inductive inferences that involve reasoning from one fairly specific statement to another, as in argument (4).

(4) Ferrets have sesamoid bones.
Squirrels have sesamoid bones.

There is now a well-documented set of empirical regularities on inductive reasoning. We provide an introduction to these empirical regularities and then describe theoretical accounts of inductive reasoning (see Heit, 2000, for a more extensive review).

Key Results in Inductive Reasoning

One of the early experimental studies of inductive reasoning, by Rips [1975], looked at how people project properties of one category of animals to another. Subjects were told to assume that on a small island, it has been discovered that all members of a particular species have a new type of contagious disease. Then subjects judged for various other species what proportion would also have the disease. For example, if all rabbits have this disease, what proportion of dogs have the disease? Rips used a variety of animal categories in the premise and conclusion roles. It was found that two factors consistently promoted inferences from a premise category to a conclusion category. First, similarity between premises and conclusions promoted strong inferences. For example, subjects made stronger inferences from rabbits to dogs than from rabbits to bears. Second, the typicality of the premise, with respect to its superordinate category, was critical in promoting inferences. The result was that more typical premise categories led to stronger inferences than atypical premise categories. For example, with the bird stimuli, having *bluejay* as a premise category led to stronger inferences overall compared to having *goose* as a premise category. Using multiple regression analyses, Rips found distinct contributions of premise-conclusion similarity and premise typicality. Interestingly, there was no evidence for a role of conclusion typicality. For example, all other things being equal, people would be as willing to draw a conclusion about a bluejay or about a goose, despite the difference in typicality of these two categories (see [Osherson *et al.*, 1990], for further investigations of similarity and typicality effects).

The next major study of induction was by Nisbett *et al.*, [1983], who also asked subjects to draw inferences about items (animals, people, and objects) found on a remote island. For example, subjects were told to imagine that one member of the Barratos tribe is observed to be obese, and they estimated the proportion of all members of this group that would be obese. Likewise, subjects were told that one sample of the substance "floridium" was observed to conduct electricity, and they estimated the proportion of all members of this set that would conduct electricity. One key finding was that subjects were very sensitive to perceived variability of the conclusion category. For a variable category such as Barratos people (and their potential obesity), subjects were rather unwilling to make strong inferences about other Barratos, after just one case. But for a non-variable category such as floridium samples, subjects were willing to generalize the observation of electrical conductance to most or all of the population. This result, that subjects are more willing to draw inferences about less variable conclusion categories, makes a striking contrast to the results of Rips [1975]. Whereas Rips found that typicality of the conclusion did not affect inductive strength, Nisbett et al. showed that conclusion categories do matter, at least in terms of their variability.

The preceding results show how people reason based on a single premise. However, when people try to make an inference about some object or event, they are typically faced with a great deal of information. Rather than just one past case being available or relevant, in many realistic situations there will be an extensive set of cases or premises that could be relied on. What makes a set of premises seem strong, or useful for promoting inferences? One factor is numerosity. In their study involving inferences about people and objects on an island, Nisbett *et al.* [1983] systematically varied the given number of observations. For example, subjects were told that 1, 3, or 20 obese members of the Barratos group had been observed, and asked what proportion of all Barratos are obese. In general, inferences were stronger with increased sample size (see also [Osherson *et al.*, 1990]).

Although sheer numerosity of cases does have some effect on induction, there is also substantial evidence that variability or diversity of cases affects inductive strength. Intuitively, repeating the same evidence, or highly similar pieces of evidence, again and again should not be much more convincing that just giving the evidence once. Consider the following arguments (adapted from [Osherson *et al.*, 1990]).

(5) Cows require vitamin K for the liver to function.
Horses require vitamin K for the liver to function.
———
All mammals require vitamin K for the liver to function.

(6) Cows require vitamin K for the liver to function. (6)
Ferrets require vitamin K for the liver to function.
———
All mammals require vitamin K for the liver to function.

Although both arguments seem to have some argument strength, most people find argument (6) to be stronger than argument (5), due to the greater diversity of premise information. Again, there is an interesting comparison to Nisbett *et al.* [1983], who found that variable conclusions led to weaker inductive inferences. In contrast, it has been found that diverse premise categories lead to stronger inductive inferences. Another fascinating aspect of the diversity effect is that it runs in the opposite direction to the typicality effect: Whereas a typical premise category leads to a fairly strong inductive argument (1), an argument with two typical premise categories (5) is actually weaker than an argument with a typical premise and an atypical premise (6).

Effects of Knowledge on Inductive Reasoning

Unlike deductive reasoning, where it should be possible to determine just from the form of an argument whether the conclusion must necessarily follow, inductive reasoning is uncertain by nature. Hence it should be rational to go beyond the information given, seeking other knowledge that could reduce this uncertainty and make inductive inferences more accurate. Indeed, all of the examples of inductive reasoning in this section rely on some use of world knowledge that is not explicitly stated in the inductive arguments, such as that cows and horses are more similar than are cows and ferrets. However, in other ways researchers have aimed to study the "essence" of inductive reasoning by discouraging the use of outside knowledge. For example, Rips [1975] used fictional diseases that people would not have strong prior beliefs about and Osherson *et al.* [1990] used "blank" properties such as "has sesamoid bones" which sounded somewhat biological but were fairly unfamiliar. These decisions by these researchers were helpful indeed in uncovering the various empirical regularities such as similarity, typicality, and diversity effects.

Still, other researchers have studied the role of knowledge in induction more directly. For example, Medin *et al.* [1997] looked at inductive reasoning about categories of plants, by various kinds of tree experts, such as taxonomists and tree maintenance workers. Here the main interest was effects of similarity, for groups that differed in their notions of similarity. For example, in a sorting task, maintenance workers tended to organize tree species in terms of their shape or purpose for various landscaping tasks. Medin et al. devised questions on a test of inductive reasoning that pitted scientific matches against alternative, functional category structures. For example, two tree species might be distant in terms of the scientific taxonomy but they could both be useful for providing shade. It was found that taxonomists (not surprisingly) sorted trees on the basis of scientific taxonomy and likewise favored inductive arguments between categories that were close in the scientific taxonomy. Maintenance workers seemed to favor a more functional category organization for both sorting and reasoning. In sum, the groups of experts generally showed the similarity effects that had been documented in other studies of induction, but their knowledge about trees mediated these similarity effects.

Other evidence for knowledge effects has highlighted the effects of the property that is being inferred. The Nisbett *et al.* [1983] study is a good illustration of how knowledge about the scope of a property affects inductive inference. As already reviewed, seeing that just one member of the Barratos group is obese does not seem to promote the inference that other people in this group will be obese. Obesity seems to be more of an individual characteristic rather than a group characteristic. On the other hand, Nisbett et al. found that people make stronger inferences for the same category but another property, skin color. Here, seeing the skin color of just one Barratos promotes inferences about other members of this group, on the assumption that members of the same ethnic group will likely have some shared physical characteristics. (See [Goodman, 1955] for further discussion of how properties differ in their tendency to promote induction.)

Although it might seem from the previous section that some properties have a wider scope for inference than others, the picture is actually more complicated. Depending on the categories in an inductive argument, a particular property may lead to strong inferences or weak inferences or something in between. Consider the following example, from [Heit and Rubinstein, 1994]. For an anatomical property, such as "has a liver with two chambers," people will make stronger inferences from chickens to hawks than from tigers to hawks. Because chickens and hawks are from the same biological category, and share many internal properties, people are quite willing to project a novel anatomical property from one bird to another. But since tigers and hawks differ in terms of many known internal biological properties, it seems less likely that a novel anatomical property will project from one to the other. However, now consider the behavioral property "prefers to feed at night." Heit and Rubinstein [1994] found that inferences for behavioral properties concerning feeding and predation were weaker between the categories *chicken* and *hawk* than between the categories *tiger* and *hawk* — the opposite of the result for anatomical properties. Here, it seems that despite the major biological differences between tigers and hawks, people were influenced by the known similarities between these two animals in terms of predatory behavior, thus making strong inferences about a novel behavioral property

Theoretical Accounts of Inductive Reasoning

So far, we have described several empirical regularities in inductive reasoning, including similarity effects, typicality effects, diversity effects, and effects based on knowledge about the property being inferred. Together, these results pose a challenge for psychological accounts of induction. Although there have been a number of proposals (see, in particular, [Osherson *et al.*, 1990; Sloman, 1993]), we will focus on a model of inductive reasoning by Heit [1998] (see also [Tenenbam and Griffith, 2001; Kemp and Teenbaum, 2009]) that has been applied to all of these results. This is a model derived from Bayesian statistics and we will show that people's inductive reasoning behaviour does indeed seem to follow the dictates of inductive logic.

According to the Bayesian model, evaluating an inductive argument is conceived of as learning about a property, in particular learning for which categories the property is true or false. For example, in argument (1) above, the goal would be to learn which animals have sesamoid bones and which animals do not. The model assumes that for a novel property such as in this example, people would rely on prior knowledge about familiar properties, to derive a set of hypotheses about what the novel property may be like. For example, 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. Is it an all-mammal property, or a cow-only property? What is crucial is that people assume that novel properties follow the same distribution as known properties. Because many known properties of cows are also true of other mammals, argument (1) regarding a novel property seems fairly strong.

The Bayesian model addresses many of the key results in inductive reasoning. For example, the model can predict similarity effects as in [Rips, 1975]. Given that rabbits have some kind of disease, it seems more plausuble to infer that dogs have the same disease rather than bears, because rabbits and dogs are more alike in terms of known properties than are rabbits and bears. The Bayesian model also addresses typicality effects, under the assumption that according to prior beliefs, atypical categories, such as *geese*, would have a number of idiosyncratic features. Hence a premise asserting a novel property about geese would suggest that this property is likewise idiosyncratic and not to be widely projected. In contrast, prior beliefs about typical categories, such as *bluejays*, would indicate that they have many properties in common with other categories, hence a novel property of a typical category should generalize well to other categories.

The Bayesian model also addresses diversity effects, with a rationale similar to that for typicality effects. An argument with two similar premise categories, such as cows and horses in (5), could bring to mind a lot of idiosyncratic properties that are true just of large farm animals. Therefore a novel property of cows and horses might seem idiosyncratic to farm animals, and not applicable to other mammals. In contrast, an argument with two diverse premise categories, such as cows and ferrets in (6), could not bring to mind familiar idiosyncratic properties that are true of just these two animals. Instead, the prior hypotheses would be derived from known properties that are true of all mammals or all animals. Hence a novel property of cows and ferrets should generalize fairly broadly.

To give a final illustration of the Bayesian approach, when reasoning about the anatomical and behavioral properties in [Heit and Rubinstein, 1994], people could draw on prior knowledge about different known properties for the two kinds of properties. Reasoning about anatomical properties could cause people to rely on prior knowledge about familiar anatomical properties. In contrast, when reasoning about a behavioural property such as "prefers to feed at night," the prior hypotheses could be drawn from knowledge about familiar behavioural properties. These two different sources of prior knowledge would lead to different patterns of inductive inferences for the two kinds of properties.

Summary: Inductive reasoning

To conclude, the Bayesian model does address a fairly broad set of phenomena (see [Heit, 1998; 2000] for further applications, in greater detail). There are other models, such as those proposed by Osherson *et al.* [1990] and Sloman [1993], that can address many of the same results, however we see a big advantage of the Bayesian model is that it derives from the same principles, probability theory and Anderson's [1990; 1991] rational analysis, as do recent models of deduction to which we now turn.

4 DEDUCTIVE REASONING

In this section, we review recent work which suggests that empirical research on putatively deductive reasoning tasks is better characterised using inductive logic. Empirical studies of deductive reasoning have concentrated on three main experimental tasks, conditional inference, data selection, and quantified syllogistic reasoning. A subsection is devoted to each task. In each, we describe recent

Bayesian probabilistic models that seem able to account for the deviations from deductive prescriptions seen in the experimental results. The key idea behind all these models is to use conditional probability, $P(q|p)$, to account for the meaning of conditional statements, *if p then q* (e.g., *if you turn the key then the car starts*). For each area of reasoning, we introduce the task, and the standard findings. We then introduce a Bayesian rational analysis for each problem, show how it accounts for the core data, and how it generalises to a sample of further important data in the area.

Conditional Inference

In conditional inference four inference patterns have been extensively studied experimentally: the valid inference forms *modus ponens* (MP) and *modus tollens* (MT) and the fallacies *denying the antecedent* (DA) and *affirming the consequent* (AC). Each inference consists of the *conditional* premise and one of four possible *categorical* premises, which relate either to the antecedent or consequent of the conditional, or their negations ($p, \neg p, q, \neg q$ where "$\neg$" = not). For example, the inference Modus Ponens (MP) combines the conditional premise *if p then q* with the categorical premise p; and yields the conclusion q.

According to standard logic, we would expect everyone to endorse the valid inferences and not to endorse the fallacies. However, people tend endorse all four inferences at rates above 50% and in a characteristic order: MP > MT > AC > DA [Schroyens and Schaeken, 2003]. All the difference in endorsement rate between pairs in the order are highly statistically significant. This performance reveals a large divergence between people's behaviour the predictions of the standard logical model.

A Probabilistic Approach

In empirical psychology, there are a variety of probabilistic approaches to conditional inference [Anderson, 1995; Liu, 2003; Evans and Over, 2004; Pfeifer and Kleiter, 2005; Oaksford and Chater, 2007; Oaksford *et al.*, 2000]. Apart from Evans and Over [2004], these approaches have attempted to explain human reasoning performance without invoking a particular psychological implementation of inductive logic. All these accounts share three key ideas. First, the probability of a conditional is the conditional probability, i.e., P (*if p then q*) $= P(q|p)$. In the normative literature, this identification is simply called "The Equation" [Adams, 1998; Bennett, 2003; Edgington, 1995]. In the psychological literature, the Equation has been confirmed experimentally by Evans, Handley, and Over [2003]; see also, [Over *et al.*, 2007] and by Oberauer and Wilhelm [2003]. Second, as discussed above, probabilities are interpreted "subjectively," that is, as degrees of belief. It is this interpretation of probability that allows us to provide a probabilistic theory of inference as belief updating. Third, conditional probabilities are determined by a psychological process called the "Ramsey Test" [Bennett, 2003; Ramsey, 1931/1990]. For example, suppose you want to evaluate your conditional degree of belief that *if it is sunny in Wimbledon*, *then John plays tennis*. By the Ramsey test, you make the hypothetical supposition that *it is sunny in Wimbledon* and revise your other beliefs so that they fit with this supposition. You then "read off" your hypothetical degree of belief that *John plays tennis* from these revised beliefs.

Liu [2003] and Oaksford *et al.* [2000]; see also, [Oaksford and Chater, 2007] treat conditional inference as belief revision. We concentrate on this approach because it seems to provide the possibility of accounting for human performance with the minimal additional assumptions about the cognitive system. Treating conditional inference as belief revision concerns how we reason when the categorical premise is not merely *supposed*, but is actually believed or known to be true. This process is known as conditionalisation. Consider an MP inference, e.g., *if it is sunny in Wimbledon*, *then John plays tennis*, and *it is sunny in Wimbledon*, therefore, *John plays tennis*. Conditionalisation applies when we know (instead of merely supposing) that *it is sunny in Wimbledon*; or when a high degree of belief can be assigned to this event (e.g., because we know that it is sunny in nearby Bloomsbury). By conditionalisation, our new degree of belief that *John plays tennis* should be equal to our prior degree of belief that *if it is sunny in Wimbledon*, *then John plays tennis* (here "prior" means before learning that it is sunny in Wimbledon). More formally, by the Equation, we know that P_0 (*if it is sunny in Wimbledon*, *then John plays tennis*) equals P_0(*John plays tennis*|*it is sunny in Wimbledon*), where "$P_0(x)$" = prior probability of x. When we learn *it is sunny in Wimbledon*, then P_1(*it is sunny in Wimbledon*) = 1, where "$P_1(x)$" = posterior probability of x. Conditionalising on this knowledge tells us that our new degree of belief in *John plays tennis* P_1(*John plays tennis*), should be equal to P_0(*John plays tennis*|*it is sunny in Wimbledon*). That is, $P_1(q) = P_0(q|p)$, where p = *it is sunny in Wimbledon*, and q = *John plays tennis*.[1] So from a probabilistic perspective, MP provides a way of updating our degrees of belief in the consequent, q, on learning that the antecedent, p, is true.

Quantitatively, if you believe that P_0(*John plays tennis*|*it is sunny in Wimbledon*) = .9, then given you discover that *it is sunny in Wimbledon* (P_1(*it is sunny in Wimbledon*) = 1) your new degree belief that *John plays tennis* should be .9, i.e., P_1(*John plays tennis*) = .9. This contrasts with the logical approach in which believing the conditional premise entails with *certainty* that the conclusion follows from the minor premise so that P_0(*John plays tennis*|*it is sunny in Wimbledon*) = 1. This is surely too strong a claim.

The extension to the other conditional inferences is not direct, however. Take an example of (AC), *if it is sunny in Wimbledon*, *John plays tennis* and *John plays tennis*, therefore, *it is sunny in Wimbledon*. In this case, one knows or strongly

[1]The case where the categorical premise is uncertain can be accommodated somewhat controversially using a generalization of this idea, Jeffrey conditionalisation [Jeffrey, 1983]. The new degree of belief that *John plays tennis* (q), on learning that *it is sunny in Bloomsbury* (which confers only a high probability that *it is sunny in Wimbledon* (p)), is:

$P_1(q) = P_0(q|p)P_1(p) + P_0(q|\neg p)P_1(\neg p)$.

believes that *John plays tennis* (perhaps we were told by a very reliable source), so $P_1(q) = 1$. But to use Bayesian conditionalisation to infer one's new degree of belief that *it is sunny in Wimbledon*, $P_1(p)$, one needs to know one's conditional degree of belief that *it is sunny in Wimbledon* given *John plays tennis*, i.e., $P_0(p|q)$. However, the conditional premise of AC, like that of MP, is about $P_0(q|p)$ *not* about $P_0(p|q)$ [Sober, 2002]. The solution proposed by Oaksford *et al.*, [2000] (see also [Wagner, 2004]) is that that people also know the prior marginal probabilities (at least approximately). That is, they know something about the probability of a sunny day in Wimbledon, $P_0(p)$, and the probability that John plays tennis, $P_0(q)$, *before* learning that it is in fact a sunny day in Wimbledon. With this additional information, $P_0(p|q)$ can be calculated from the converse conditional probability, $P_0(q|p)$, using Bayes' Theorem.[2] The same approach also works for the two othr types of conditional inferene, Denying the Antecedent (DA) and Affirming the Consequent (AC) where the relevant probabilities are $P_0(\neg q|\neg p)$ and $P_0(\neg p|\neg q)$ respectively. The fact that the conditional premises of AC, DA and MT do not determine the appropriate conditional probability marks an important asymmetry with MP. For these inferences, further knowledge is required to infer the relevant conditional degrees of belief.

The Empirical Data

We now show how some of the errors and biases observed in conditional inference can be seen as a consequence of this rational probabilistic model. The first set of "biases" are called "the inferential asymmetries" [Oaksford and Chater, 2008]. That is, MP is drawn more than MT and AC is drawn more than DA (MT is also drawn more than AC). Oaksford and Chater [2003; 2007; 2008] calculated the values of $P_0(q|p)$, $P_0(p)$ and $P_0(q)$ that best fit the data, i.e., they minimize the sum of squared error between the data and the models predictions. The fits were good ($R^2 = .84$) and the probabilities, $P_0(q|p) = .88$, $P_0(p) = .54$, and $P_0(q) = .70$, seems reasonable, i.e., $P_0(q|p)$ is high, $P_0(q|p) \approx .5$, and $P_0(q) > P_0(p)$. To predict John's tennis playing behaviour well $P_0(q|p)$ should be high. Further, one would be unlikely to draw inferences about John's tennis playing behaviour using this rule in contexts where the probability that it was sunny was less than chance [Adams, 1998]. Moreover, as long as $P_0(q|p)$ high $P_0(q) > P_0(p)$ is most likely to hold. However, this probabilistic model [Oaksford *et al.*, 2000] does not capture the magnitudes of the inferential asymmetries [Evans and Over, 2004; Schroyens and Schaeken, 2003]. It underestimates the MP–MT asymmetry and overestimates the DA–AC asymmetry.

Oaksford and Chater [2007] argued that this is because learning that the categorical premise is true can have two inferential roles. The first inferential role is in conditionalisation, as we have described. The second inferential role is based on

[2]Bayes' theorem is the elementary identity of probability theory mentioned above that allows a conditional probability to be calculated from its converse conditional probability and the priors: $P(p|q) = (P(q|p)P(p))/P(q)$.

the pragmatic inference that *being told that the categorical premise is true* often suggests that there is a counterexample to the conditional premise. For example, consider the MT inference on the rule *if I turn the key the car starts*. If you were told that *the car did not start*, it seems unlikely that you would immediately infer that the *key was not turned*. Telling someone that *the car did not start* seems to presuppose that an attempt has been made to start it, presumably by turning the key. Consequently, the categorical premise here seems to suggest a counterexample to the conditional itself, i.e., a case where the key was turned but the car did not start. Hence one's degree of belief in the conditional should be reduced on *being told* that the car did not start. Notice, here, the contrast between being told that the car did not start (and drawing appropriate pragmatic inferences), and merely *observing* a car that has not started (e.g., a car parked in the driveway). In this latter situation, it is entirely natural to use the conditional rule to infer that the key has not been turned.

Where the second, pragmatic, inferential role of the categorical premise is operative, this violates what is called the *rigidity condition* on conditionalisation, $P_0(q|p) = P_1(q|p)$ [Jeffrey, 1983]. That is, learning the categorical premise alters ones degree of belief in the conditional premise. Oaksford and Chater [2007; 2008] argue that taking account of such rigidity violations helps capture the probability of the conditional; and that, for MT, this modified probability is then used in conditionalisation. Furthermore, they argue that DA and AC also suggest violations of the rigidity condition, concerning the case where the car starts without turning the key. These violations lead to reductions in ones's degree of belief that the cars starts, given that the key is turned ($P_0(q|p)$). Using this lower estimate to calculate the relevant probabilities for DA, AC and MT can rationally explain the relative magnitudes of the MP–MT and DA–AC asymmetries (see Figure 2, Panel D).

Another one of the key empirical biases of conditional inference is *negative conclusion* bias. This bias arises when negations are used in conditional statements, e.g., *if a bird is a swan, then it is not red*. In Evans' [1972] *Negations Paradigm*, four such rules are used, *if p then q, if p then not-q, if not-p then q, and if not-p then not-q*. The most robust finding is that people endorse DA, AC, and MT more when the conclusion contains a negation. So, for example, DA on *if p then q* yields a negated conclusion, *not-q*, whereas, DA on *if p then not-q* yields an affirmative conclusion, q (because *not-not-q* $= q$). In the data, the frequency with which DA is endorsed for *if p then q* is much higher than for *if p then not-q*.

To explain negative conclusion bias, Oaksford *et al.* [2000] appealed to the idea that most categories apply only to a minority of objects [Oaksford and Stenning, 1992]. Hence, the probability of an object being, say, red is lower than the probability of it not being red, i.e., $P_0(Red) < P_0(\neg Red)$. Consequently, the marginal probabilities ($P_0(p)$ and $P_0(q)$) will take on higher values when p or q are negated. Higher values of the prior probabilities of the conclusion imply higher values of the relevant conditional probabilities for DA, AC and MT, i.e., to higher values of the posterior probability of the conclusion. So, for example, for our rule *if* *a bird is a swan, then it is white*, the prior probability of the conclusion of the DA inference ($P_0(\neg \mathit{White})$) is high. This means that the conditional probability ($P_0(\neg \mathit{White}|\neg \mathit{Swan})$) is also high and, consequently, so is the probability of the conclusion ($P_1(\neg \mathit{White})$). Therefore, an apparently irrational negative conclusion bias can be seen as a rational "high probability conclusion" effect. Oaksford *et al.* [2000] tested this explanation by manipulating $P_0(p)$ and $P_0(q)$ directly rather than using negations and showed results closely analogous to negative conclusion bias.

To conclude this section on conditional inference, we briefly review one of the most cited problems for a probabilistic account. Like any computational level analysis, this account avoids theorising about the specific mental representations or algorithms involved in conditional reasoning. This may seem unsatisfactory. We suggest, by contrast, that it is premature to attempt an algorithmic analysis. The core of the probabilistic approach interprets conditionals in terms of conditional probability, i.e., using the Equation; and our current best understanding of conditional probability is given by the Ramsey test [Bennett, 2003]. But there is currently no possibility of building a full algorithmic model to carry through the Ramsey test, because this involves solving the notorious frame problem [Pylyshyn, 1987]. That is, it involves knowing how to update one's knowledge-base, in the light of a new piece of information — and this problem has defied 40 years of artificial intelligence research.

Nonetheless, an illustrative small-scale implementation of the Ramsey test is provided by the operation of a constraint satisfaction neural network [Oaksford, 2004; Oaksford and Chater, in press]. In such a model, performing a Ramsey test means clamping on or off the nodes or neurons corresponding to the categorical premise of a conditional inference. Network connectivity determines relevance relations and the weight matrix encodes prior knowledge. Under appropriate constraints, such a network can be interpreted as computing true posterior probabilities [McClelland, 1998]. A challenge for the future is to see whether such small-scale implementations can capture the full range of empirically observed effects in conditional inference.

Data Selection

Data selection involves choosing data to confirm or disconfirm a hypothesis, and it has been extensively investigated empirically using Wason's [1968] selection task. This task has featured prominently in the philosophical discussions about human rationality (e.g., [Cohen, 1980; Stich, 1985; Stein, 1996]. In this task, people see four double-sided cards, with a number on one side and a letter on the other. They are asked which cards they should turn over, in order to test the hypothesis that *if there is an A (p) on one side of a card, then there is a 2 (q) on the other.* The upturned faces of the four cards show an A (p), a K ($\neg p$), a 2 (q), and a 7 ($\neg q$). As Popper [1959/1935] argued, logically one can never be *certain* that a scientific hypothesis is true in the light of observed evidence, as the very next piece of evidence one discovers could be a counterexample. So just because all the swans you have observed up until now have been white is no guarantee that the next one will not be black. Instead, Popper argues that the only logically sanctioned strategy for hypothesis testing is to seek *falsifying* cases. In testing a conditional rule *if p then q*, this means seeking out p, $\neg q$ cases. This means that, in the standard selection task, one should select the A (p) and the 7 ($\neg q$) cards, because these are the only cards that could potentially falsify the hypothesis. However, as for conditional inference, there is a large divergence between this logical prediction and the data. Indeed, rather than seek *falsifying* evidence, participants seem to select the cases that *confirm* the conditional, i.e., the A (p) and the 2 (q). This is called "confirmation bias."

A Probabilistic Approach

As with conditional inference, a variety of probabilistic approaches to data selection have been proposed [Evans and Over, 1996a; 1996b; Klauer, 1999; Nickerson, 1996; Over and Evans, 1994, Over and Jessop, 1998], which they all originate from the optimal data selection (ODS) model of Oaksford and Chater [1994] (see also, [1996; 2003b]). This model is derived from the normative literature on optimal experimental design in Bayesian statistics [Lindley, 1956]. The idea again relies on interpreting a conditional in terms of conditional probability. For example, the hypothesis, *if swan* (p) *then white* (p), is interpreted as making the claim that the probability of a bird being white given that it is a swan, $P(q|p)$, is high, certainly higher than the base rate of being a white bird, $P(q)$. This hypothesis is called the *dependence* hypothesis (H_D). Bayesian hypothesis testing is comparative rather than exclusively concentrating on falsification. Specifically, in the ODS model, it is assumed that people compare H_D with an *independence* hypothesis (H_I) in which the probability of a bird being white, given it is a swan, is the same as the base rate of a bird being white, i.e., $P(q|p) = P(q)$. We assume that, initially, people are maximally uncertain about which hypothesis is true ($P(H_D) = P(H_I)$ $= 0.5$) and that their goal in selecting cards is to reduce this uncertainty as much as possible while turning the fewest cards.

Take, for example, the card showing *swan* (p). This card could show *white* on the other side (p, q) or another color ($p, \neg q$). The probabilities of each outcome will be quite different according to the two hypotheses. For example, suppose that the probability of a bird being white, given that it is a swan is .9 ($P(q|p, \mathrm{H}_D) = .9$) in the dependence hypothesis; the marginal probability that a bird is swan is .2 ($P(p)$ $= .2$); and the marginal probability that a bird is white is .3 ($P(q) = .3$). Then, according to the dependence hypothesis, the probability of finding *white* (q) on the other side of the card is .9, whereas according to the independence hypothesis it is .3 (as the antecedent and consequent are, in this model, independent, we need merely consult the relevant marginal probability). And, according to the dependence hypothesis, the probability of finding a colour other than white ($\neg q$) on the other side of the card is .1, whereas, according to the independence hypothesis, it is .7. With this information, it is now possible to calculate one's new degree of uncertainty about the dependence hypothesis after turning the *swan* card to find *white* on the other side ($P(\mathrm{H}_D|p, q)$). According to Bayes' theorem, this probability is .75. Hence, one's new degree of belief in the dependence model should be .75 and one's degree of belief in the independence model should be .25. Hence, the degree of uncertainty about which hypothesis is true has been reduced. More specifically, the ODS model is based on *information gain*, where information is measured in *bits* as in standard communication theory [Shannon and Weaver, 1949]. Here, the initial uncertainty is 1 bit (because $P(\mathrm{H}_D) = P(\mathrm{H}_I) = 0.5$, equivalent to the uncertainty of a single fair coin flip) and in this example this is reduced to .81 bits (because now $P(\mathrm{H}_D) = .75$ and $P(\mathrm{H}_I) = 0.25$). This is an *information gain* of .19 bits.

In Wason's task, though, participants do not actually turn the cards, and hence they cannot know how much information they will gain by turning a card before doing so. Consequently, they must base their decision on *expected* information gain, taking both possible outcomes (p, q and $p, \neg q$) into account. The ODS model assumes that people select each card in direct proportion to its expected information gain.

The ODS model also makes a key assumption about the task environment the *rarity* assumption: that the properties that occur in the antecedents and consequents of hypotheses are almost always rare and so have a low base rate of occurrence. For example, most birds are not swans and most birds are not white. That people make this assumption has received extensive independent verification [McKenzie *et al.*, 2001; McKenzie and Mikkelsen, 2000; 2007].

The Empirical Data

The ODS model predicts that the two cards that lead to the greatest expected information gain are the p and the q cards. Fitting the model to the data, reveals a good fit [Oaksford and Chater, 2003b] and when $P(q|p, \mathrm{H}_D)$ was set to .9 the best fitting values of $P(p)$ and $P(q)$ were .22 and .27 respectively, i.e., very close to the values used in the above example. The ODS model suggests that performance on the selection task displays rational hypothesis testing behaviour, rather than irrational confirmation bias. Taking rarity to an extreme provides a simple intuition here. Suppose we consider the (rather implausible) conditional *if a person is bitten by a vampire bat* (p)*, they will develop pointed teeth* (q). Clearly, we should check people who we know to have been bitten, to see if their teeth are pointed (i.e., turn the p card); and, uncontroversially, we can learn little from people we know have not been bitten (i.e., do not turn the $\neg p$ card). If we see someone with pointed teeth, it is surely worth finding out whether they have been bitten — if they have, this raises our belief in the conditional, according to a Bayesian analysis (this is equivalent to turning the q card). But it seems scarcely productive to investigate someone without pointed teeth (i.e., do not turn the $\neg q$ card) to see if they have been bitten. To be sure, it is *possible* that such a person might have been bitten, which would disconfirm our hypothesis, and lead to maximum information gain; but this has an almost infinitesimal probability. Almost certainly, we shall find that they have not been bitten, and learn nothing. Hence, with rarity, the expected informativeness of the q card is higher than that of the $\neg q$ card, diverging sharply from the falsificationist perspective, but agreeing with the empirical data.

It has been suggested, however, that behaviour on this task might be governed by what appears to be a wholly non-rational strategy: *matching bias*. This bias arises in the same context as negative conclusion bias that we discussed above, i.e., in Evans' [1972] negations paradigm. Take, for example, the rule *if there is an A on one side*, *then there is not a 2 on the other side* (*if p then* $\neg q$).The cards in this task are described using their logical status, so for this rule, 2 is the *false consequent* (FC) card and 7 is the *true consequent card* (TC). For this negated consequent rule, participants tend to select the A card (TA: *true antecedent*) and the 2 card (FC). That is, participants now seem to make the falsifying response. However, as Evans [1972] pointed out, participants may simply ignore the negations entirely and *match* the values named in the conditional, i.e., A and 2. *Prima facie*, this is completely irrational. However, the "contrast set" account of negation shows that due to the rarity assumption — that most categories apply to a minority of items — *negated* categories are high probability categories (see above). Having a high probability antecedent or consequent alters the expected information gains associated with the cards. If the probability of the consequent is high then the ODS model predicts that people should make the falsifying TA and FC responses, because these are associated with the highest information gain. Consequently, matching bias is a rational hypothesis testing strategy after all.

Probabilistic effects were first experimentally demonstrated using the *reduced array* version of Wason's selection task [Oaksford *et al.*, 1997], where participants can successively select up to 15 q and 15 $\neg q$ cards (there are no upturned p and $\neg p$ cards that can be chosen). As predicted by the ODS model, where the probability of q is high (i.e., where rarity is violated), participants select more $\neg q$ cards and fewer q cards. Other experiments have also revealed similar probabilistic effects [Green and Over, 1997; 2000; Kirby, 1994; Oaksford *et al.*, 1999; Over and Jessop, 1998].

There have also been some failures to produce probabilistic effects, however (e.g., [Oberauer *et al.*, 1999; 2004]). It has been argued that these arise because of weak probability manipulations or other procedural problems [Oaksford and Chater, 2003b; Oaksford and Moussakowski, 2004; Oaksford and Wakefield, 2003]). Using a *natural sampling* [Gigerenzer and Hoffrage, 1995] procedure, in which participants sample the frequencies of the card categories while performing a selection task, probabilistic effects have been observed using using the *same* materials as Oberauer *et al.* [1999], where these effects were not evident [Oaksford and Wakefield, 2003].

In further work on matching bias, Yama [2001] devised a crucial experiment to contrast the matching bias and the information gain accounts. He used rules that introduced a high and a low probability category, relating to the blood types Rhesus Negative (Rh-) and Positive (Rh+). People were told that one of these categories, Rh-, was rare. Therefore, according to the ODS model, the rule *if p then* $\neg Rh+$ should lead participants to select the rare Rh- card. In contrast, according to matching bias they should select the Rh+ card. Yama's [2001] data were largely consistent with the information gain model. Moreover, this finding was strongly confirmed by using the natural sampling procedure with these materials [Oaksford and Moussakowski, 2004].

Alternative probabilistic accounts of the selection task have also been proposed [Evans and Over, 1996a; 1996b; Klauer, 1999; Nickerson, 1996; Over and Evans, 1994, Over and Jessop, 1998]. Recently, Nelson [2005] directly tested the measures of information underpinning these models, including Bayesian diagnosticity [Over and Evans, 1994; Evans and Over, 1996b; McKenzie and Mikkelsen, 2007], information gain [Oaksford and Chater, 1994; 1996; 2003b; Hattori, 2002], Kullback-Liebler distance [Klauer, 1999; Oaksford and Chater, 1996], probability gain (error minimization) (Baron, 1981, 1985), and impact (absolute change) [Nickerson, 1996]. Using a related data selection task, he looked at a range of cases where these norms predicted different orderings of informativeness, for various data types. Nelson found the strongest correlations between his data and information gain (.78). Correlations with diagnosticity (-.22) and log diagnosticity (-.41) were actually *negative*. These results mirrored Oaksford, Chater, and Grainger's [1999] results in the Wason selection task. Nelson's work provides strong convergent evidence for information gain as the index that most successfully captures people's intuitions about the relative importance of evidence.

Quantified Syllogistic Reasoning

Quantified syllogistic reasoning relates two quantified premises. Logic defines four types of quantified premise: *All*, *Some*, *Some... not*, and *None*. An example of a logically valid syllogistic argument is:

	Some Londoners (P) *are soldiers* (Q)
	All soldiers (Q) *are well fed* (R)
Therefore	*Some Londoners* (P) *are well fed* (R)

In this example, P and R are the *end terms* and Q is the *middle term* which is common to both premises. In the premises, these terms can only appear in four possible configurations which are called *figures*. When one of these terms appears before the copula verb ("are") it is called the *subject* term (in the example, P and Q) and when one appears after this verb it is called the *predicate* term (Q and R). As the premises can appear in either order there are 16 combinations and as each can be in one of four figures there 64 different syllogisms.

There are 22 logically valid syllogisms. If people are reasoning logically, they should endorse these syllogisms and reject the rest. However, observed behaviour is graded, across both valid and invalid syllogisms; and some invalid syllogisms are endorsed more than some valid syllogisms. Table 1 shows the graded behaviour over the 22 logically valid syllogisms. There are natural breaks dividing the valid syllogisms into three main groups. Those above the single line are endorsed most, those below the double line are endorsed least, and those in between are endorsed at an intermediate level.

Table 1. Meta-analysis of the logically valid syllogisms showing the form of the conclusion, the number of mental models an alternative non-probabilistic psychological account [Johnson-Laird, 1983] needed to reach that conclusion, and the percentage of times the valid conclusion was drawn, in each of the five experiments analysed by Chater and Oaksford [1999].

Syllogism	Conc.	MMs	Mean
All(Q,P), *All*(R,Q)	All	1	89.87
All(P,Q), *All*(Q,R)	All	1	75.32
All(Q,P), *Some*(R,Q)	Some	1	86.71
Some(Q,P), *All*(Q,R)	Some	1	87.97
All(Q,P), *Some*(Q,R)	Some	1	88.61
Some(P,Q), *All*(Q,R)	Some	1	86.71
No(Q,P), All(R,Q)	*No*	1	92.41
All(P,Q), No(R,Q)	*No*	1	84.81
No(P,Q), All(R,Q)	*No*	1	88.61
All(P,Q), No(Q,R)	*No*	1	91.14
All(P,Q), Some...not(R,Q)	*Some...not*	2	67.09
Some...not(P,Q), All(R,Q)	*Some...not*	2	56.33
All(Q,P), Some...not(Q,R)	*Some...not*	2	66.46
Some...not(Q,P), All(Q,R)	*Some...not*	2	68.99
Some(Q,P), No(R,Q)	*Some...not*	3	16.46
No(Q,P), Some(R,Q)	*Some...not*	3	66.46
Some(P,Q), No(R,Q)	*Some...not*	3	30.38
No(P,Q), Some(R,Q)	*Some...not*	3	51.90
Some(Q,P), No(Q,R)	*Some...not*	3	32.91
No(Q,P), Some(Q,R)	*Some...not*	3	48.10
Some(P,Q), No(Q,R)	*Some...not*	3	44.30
No(P,Q), Some(Q,R)	*Some...not*	3	26.56

Note The means in the final column are weighted by sample size.

A Probabilistic Approach

There has only been one probabilistic approach developed for syllogisms. This is the *Probability Heuristics Model* (PHM), [Chater and Oaksford, 1999], which was developed at both the computational and the algorithmic levels. One of the primary motivations for this model was the hypothesis that, from a probabilistic point of view, reasoning about *all* and *some* might be continuous with reasoning about more transparently probabilistic quantifiers, such as *most* and *few*. By contrast, from a logical stand point, such *generalised quantifiers* require a different, and far more complex, treatment [Barwise and Cooper, 1983], far beyond the resources of existing logic-based accounts in psychology. Perhaps for this reason, although generalized quantifiers were discussed in early mental models theory [Johnson-Laird, 1983], no empirical work on these quantifiers was carried out in the psychology of reasoning.

In deriving PHM, the central first step is to assign probabilistic meanings to the central terms of quantified reasoning using conditional probability. Take the universally quantified statement, *All P are Q* (we use capitals to denote predicates; these should be applied to variables x which are bound by the quantifier, e.g., $P(x)$, but we usually leave this implicit). Intuitively, the claim that *All Londoners are soldiers* can naturally be cast in probabilistic terms: as asserting that the probability that a person is a soldier given that they are a Londoner is 1. More generally, the probabilistic interpretation of *All* is straightforward: because its underlying logical form can be viewed as a conditional, i.e., $All(x)(if\ P(x)\ then\ Q(x))$. Thus, the meaning is given as $P(Q|P) = 1$, as specifying the conditional probability of the predicate term (Q), given the subject term (P).

Similar constraints can be imposed on this conditional probability to capture the meanings of the other logical quantifiers. So, *Some P are Q* means that $P(Q|P) > 0$; *Some P are not Q* means that $P(Q|P) < 1$; and *No P are Q* means that $P(Q|P) = 0$. Thus, for example, "*Some Londoners are soldiers*" is presumed to mean that the probability that a person is a soldier given that they are a Londoner is greater than zero, and similarly for the other quantifiers. Such an account generalises smoothly to the generalised quantifiers *most* and *few*. *Most P are Q* means that $1 - \Delta < P(Q|P) < 1$ and *Few P are Q* means that $0 < P(Q|P) < \Delta$, where Δ is small. So, for example, *Most Londoners are soldiers* may be viewed as stating that the probability that a person is a soldier, given that they are a Londoner is greater than, say, .8, but less than 1.

At the computational level, these interpretations are used to build very simple graphical models (e.g., [Pearl, 1988]) of quantified premises, to see if they impose constraints on the conclusion probability. For example, take the syllogism:

$$\begin{array}{lll} & \textit{Some P are Q} & \\ & \textit{All Q are R} & P \rightarrow Q \rightarrow R \\ \textit{Therefore} & \textit{Some P are R} & \end{array}$$

The syllogistic premises on the left define the dependencies on the right because of their *figure*, i.e., the arrangement of the middle term (Q) and the end terms

(P and R) in the premises. There are four different arrangements or *figures*. The different figures lead to different dependencies, with different graphical structures. Note that these dependency models all imply that the end terms (P and R) are *conditionally independent* given the middle term because there is no arrow linking P and R, except via the middle term Q. Assuming conditional independence as a default is a further assumption about the environment, an assumption not made in, for example, Adams' [1998] probability logic.

These dependency models can be parameterised. Two of the parameters will always be the conditional probabilities associated with the premises. One can then deduce whether the constraints on these probabilities, implied by the above interpretations, impose constraints on the possible conclusion probabilities, i.e., $P(R|P)$ or $P(P|R)$. In this example, the constraints that $P(Q|P) > 0$, and $P(R|Q) = 1$ and the conditional independence assumption *entail* that $P(R|P) > 0$. Consequently, the inference to the conclusion *Some P are R* is probabilistically valid (p-valid). If each of the two possible conclusion probabilities, $P(R|P)$ or $P(P|R)$, can fall anywhere in the [0, 1] interval given the constraints on the premises, then no p-valid conclusion follows. It is then a matter of routine probability to determine which inferences are $p-$valid, of the 144 two premise syllogisms that arise from combining *most* and *few* and the four logical quantifiers [Chater and Oaksford, 1999].

In PHM, however, this rational analysis is also supplemented by an algorithmic account. It is assumed that people approximate the dictates of this rational analysis by using simple heuristics. Before introducing these heuristics, though, we must introduce two key notions: the notions of the *informativeness* of a quantified claim, and the notion of *probabilistic entailment* between quantified statements.

According to communication theory, a claim is informative in proportion to how surprising it is: informativeness varies inversely with probability. But what is the probability of an arbitrary quantified claim? To make sense of this idea, we begin by making a rarity assumption, as in our models of the conditional reasoning, and the selection task, i.e., the subject and predicate terms apply to only small subsets of objects. On this assumption, if we selected subject term P, and predicate term, Q, at random, then it is very likely that they will not cross-classify any object (this is especially true, given the hierarchical character of classification, Rosch, 1975). Consequently, $P(Q|P) = 0$ and so *No P are Q* is very likely to be true, e.g., *No toupees are tables*. Indeed, for any two randomly chosen subject and predicate terms it is probable that *No P are Q*. Such a statement is therefore quite uninformative. *Some P are not Q* is even more likely to be true, and hence still less informative, because the probability interval it covers includes that for *No P are Q*. The quantified claim least likely to be true is *All P are Q*, which is therefore the most informative. Overall the quantifiers have the following order in informativeness: I(*All*) > I(*Most*) > I(*Few*) > I(*Some*) > I(*None*) > I(*Some-not*) (see [Oaksford *et al.*, 2002] for further analysis and discussion).

Informativeness applies to individual quantified propositions. The second background idea, probabilistic entailment, concerns inferential relations *between* quantified propositions. Specifically, the use of one quantifier frequently provides evidence that another quantifier could also have been used. Thus, the claims that *All swans are white* is strong evidence that *Some swans are white* — because $P(white|swan) = 1$ is included in the interval $P(white|swan) > 0$ (according to standard logic, this does not follow logically, as there may be no swans). Thus, we say that *All* probabilistically entails (or p-entails) *Some*. Similarly, *Some* and *Some. . . not* are mutually p-entailing because the probability intervals $P(Q|P) > 0$ and $P(Q|P) < 1$ overlap almost completely.

With this background in place, we can now state the probabilistic heuristics model (PHM) for syllogistic reasoning. There are two types of heuristic: *generate* heuristics which produce candidate conclusions, and *test* heuristics, which evaluate the plausibility of the candidate conclusions. The PHM account also admits the possibility that putative conclusions may also be tested by more analytic test procedures. The generate heuristics are:

(G1) *Min*-heuristic: The conclusion quantifier is the same as that of the least informative premise (*min*-premise)

(G2) *P-entailments*: The next most preferred conclusion quantifier will be the p-entailment of the *min*-conclusion

(G3) *Attachment*-heuristic: If just one possible subject noun phrase (e.g., *Some R*) matches the subject noun phrase of just one premise, then the conclusion has that subject noun phrase.

The two test heuristics are:

(T1) *Max*-heuristic: Be confident in the conclusion generated by G1–G3 in proportion to the informativeness of the most informative premise (*max*-premise)

(T2) *Some-heuristic* Avoid producing or accepting *Some_not* conclusions, because they are so uninformative.

We show how the heuristics combine in the example below:

	All P are Q	(*max*-premise)
	Some R are not Q	(*min*-premise)
Therefore	*Some_not*	(by *min*-heuristic)
	Some R are not P	(by *attachment*-heuristic)

and a further conclusion can be drawn:

Some R are P [by *p-entailment*]

Comparing the results of these heuristics with probabilistic validity, it can be shown that where there is a p-valid conclusion, the heuristics generally identify it. For example, the idea behind the *min*-heuristic is to identify the most informative conclusion that validly follows from the premises. Out of the 69 p-valid syllogisms, the *min*-heuristic identifies that conclusion for 54; for 14 syllogisms the *p*-valid conclusion is less informative than the *min*-conclusion. There is only one violation, where the *p*-valid conclusion is more informative than the *min*-conclusion.

The Empirical Data

In turning to the experimental results, we first show how all the major distinctions between standard syllogisms captured by other theories are also captured by PHM. So, returning to Table 1, all the syllogisms above the double line have the most informative *max*-premise, *All* (see heuristic T1). Moreover, all the syllogisms below the single line have uninformative conclusions *Some... not* (see heuristic T2) and those below the double line violate the *min*-heuristic (heuristic G1) and require a $p-$entailment (heuristic G2), i.e., *Some... not* $\leftrightarrow$ *Some*. Consequently, this simple set of probabilistic heuristics makes the same distinctions among the valid syllogisms as the mental models account perhaps the most influential account of syllogistic reasoning [Johson-Laird, 1983].

In this review, we concentrate on novel predictions that allow us to put clear water between PHM and other theories. As we discussed above, the most important feature of PHM is the extension to *generalised quantifiers*, like *most* and *few*. No other theory of reasoning has been applied to syllogistic reasoning with generalised quantifiers. Table 2 shows the *p*-valid syllogisms involving generalised quantifiers showing the conclusion type and the percentage of participants selecting that conclusion type in Chater and Oaksford's [1999] Experiments 1 and 2. The single lines divide syllogisms with different *max*-premises, showing a clear ordering in levels of endorsements dependent on heuristic T1. All those above the double line conform to the *min*-heuristic (heuristic G1), whereas those below do not and require a *p*-entailment (heuristic G2). As Chater and Oaksford [1999] pointed out, one difference with experiments using standard logical quantifiers was that the *Some... not* conclusion was not judged to be as uninformative, i.e., heuristic T2 was not as frequently in evidence. However, in general, in experiments using generalised quantifiers in syllogistic arguments the heuristics of PHM predict the findings just as well as for the logical quantifiers [Chater and Oaksford, 1999].

Many further results have emerged that confirm PHM. We discuss briefly discuss three of these results. First, the *min*-heuristic captures an important novel distinction between strong and weak possible conclusions introduced by Evans, Handley, Harper and Johnson-Laird [1999]. They distinguished conclusions that are necessarily true, possibly true or impossible. For example, taking the syllogism discussed earlier (with premises, *Some P are Q, All Q are R*), the conclusion *Some P are R* follows *necessarily*, *No P are R* is *impossible*, and *Some P are not R* is *possible*. Some possible conclusions are endorsed by as many participants as the necessary conclusions [Evans, *et al.* [1999]. Moreover, some of the possible conclusions were endorsed by as few participants as the impossible conclusions. Evans *et al.* [1999] observe that possible conclusions that are commonly endorsed all conform to the *min*-heuristic, whereas those which are rarely endorsed violate the

Table 2. The *p*-valid syllogisms less the syllogisms that are also logically valid (shown in Table 1), showing the form of the conclusion and the proportion of participants picking the *p*-valid conclusion in Chater and Oaksford's [1999] Experiments 1 and 2.

Syllogism	Conc.	Mean
All(Q,P), *Most*(R,Q)	Most	85
Most(Q,P), *All*(R,Q)	Most	65
All(P,Q), *Most*(Q,R)	Most	70
Most(P,Q), *All*(Q,R)	Most	55
Few(P,Q), *All*(R,Q)	Few	80
All(P,Q), *Few*(R,Q)	Few	85
Few(P,Q), *All*(R,Q)	Few	85
All(P,Q), *Few*(Q,R)	Few	75
Most(Q,P), *Most*(R,Q)	*Most*	65
Most(P,Q), *Most*(Q,R)	*Most*	50
Few(Q,R), *Most*(R,Q)	*Few*	60
Most(Q,R), *Few*(R,Q)	*Few*	75
Most(P,Q), *Few*(Q,R)	*Few*	70
Most(Q,P), *Some...not*(R,Q)	*Some...not*	80
Some...not(Q,P), *Most*(R,Q)	*Some...not*	60
Some...not(Q,P), *Most*(Q,R)	*Some...not*	75
Most(Q,P), *Some...not*(Q,R)	*Some...not*	65
Most(P,Q), *Some...not*(Q,R)	*Some...not*	75
Some...not(P,Q), *Most*(Q,R)	*Some...not*	75
Few(Q,P), *Some...not*(R,Q)	*Some...not*	60
Some...not(Q,P), *Few*(R,Q)	*Some...not*	40
Some...not(Q,P), *Few*(Q,R)	*Some...not*	30
Few(Q,P), *Some...not*(Q,R)	*Some...not*	60
Few(P,Q), *Some...not*(Q,R)	*Some...not*	60
Some...not(P,Q), *Few*(Q,R)	*Some...not*	40
All(P,Q), Most(R,Q)	*Some...not*	35
Most(P,Q), All(R,Q)	*Some...not*	35
Few(Q,P), *Few*(R,Q)	*Some...not*	35
Few(P,Q), *Few*(Q,R)	*Some...not*	30
Few(P,Q), *Most*(Q,R)	*Some...not*	30

Note This table excludes the eight MI, IM, FI, and IF syllogisms which have two *p*-valid conclusions only one of which was available in Chater and Oaksford's [1999] Experiment 2.

min-heuristic (with one exception). Hence, PHM captures this important finding.

Second, recent work relating memory span measures to syllogistic reasoning has also confirmed PHM [Copeland and Radvanksy, 2004]. PHM makes similar predictions to mental models theory because the number of heuristics that need to be applied mirrors the one, two and three model syllogism distinction (see Table 1). For one model syllogisms just the *min*-heuristic and *attachment* is required (two heuristics). For two model syllogisms, the *some...not*-heuristic is also required (three heuristics). In addition, for three model syllogisms a*p*-entailment is required (four heuristics). The more mental operations that need to be performed, the more complex the inference will be and the more working memory it will require. Copeland and Radvansky [2004] found significant correlations between working memory span and strategy use, for both mental models and PHM. While not discriminating between theories, this work confirmed the independent predictions of each theory for the complexity of syllogistic reasoning and its relation to working memory span.

Third, Copeland [2006] has provided detailed model fits to experimental data on "extended" syllogisms, i.e., syllogisms involving three quantified premises (he used only the four logical quantifiers). He fitted three different psychological models (see [Rips, 1994; Johnson-Laird and Byrne, 1991]) to these data including PHM. Using a measure of fit that penalised for complexity, he found that PHM provided better fits to the data across two experiments. This is impressive as these data only involved the logical quantifiers, that these other theories were explicitly designed to explain.

Summary: Deductive reasoning

To conclude, a Bayesian probabilistic approach to the psychology of deductive reasoning seems to make sense of a fairly broad set of phenomena that would otherwise appear to question human rationality. There are other models, such as those proposed by Rips [1994] and Johnson-Laird and Byrne [1991; 2007], that address many of the same results. However, these theories invariably deal with deviations from rationality at the algorithmic level.

5 DECISION MAKING

Whereas reasoning concerns how people use given information to derive new information, the study of decision making concerns how people's beliefs and values determine their choices. In the context of reasoning, there is fundamental debate concerning the most basic elements of a normative framework against which human performance should be compared (e.g., whether the framework should be logical [e.g., Johnson-Laird and Byrne, 1991; Rips, 1994] or probabilistic [Oaksford and Chater, 2007]). By contrast, expected utility theory is fairly widely assumed to be the appropriate normative theory to determine how, in principle, people ought to make decisions.

Expected utility theory works by assuming that each outcome, i, of a choice can be assigned a probability, $\Pr(i)$ and a utility, $U(i)$ and that the utility of an uncertain choice (e.g., a lottery ticket; or more generally, any action whose consequences are uncertain), is:

$$\sum_{\Pr}(i)U(i)$$

Expected utility theory recommends the choice with the maximum expected utility.

This normative account is breathtakingly simple, but hides what may be enormous practical complexities — both in estimating probabilities; and establishing what people's utilities are. Thus, when faced with a practical personal decision (e.g., whether to take a new job, which house to buy, whether or whom to marry), decision theory is not easy to apply — because the possible consequences of each choice are extremely complex, their probabilities ill-defined, and moreover, we often have little idea what preferences we have, even if the outcomes were definite (e.g., [Gigerenzer, 2002]). Thus, one difficulty with expected utility theory is practicability in relation to many real-world decisions. Nonetheless, where probabilities and utilities can be estimated with reasonable accuracy, expected utility is a powerful normative framework.

How far can expected utility theory be used as an explanation not merely for how agents *should* behave, but of how agents actually *do* behave? Rational choice theory, which provides a foundation for explanation in microeconomics and sociology (e.g., [Becker, 1976; 1996; Elster, 1986]) as well as perception and motor control [Körding and Wolpert, 2006], animal learning [Courville *et al.*, 2006] and behavioral ecology [Krebs and Davies, 1996; Stephens and Krebs, 1986], assumes that it does. This style of explanation involves inferring the probabilities and utilities that agents possess; and using expected utility theory to infer their choices according to those probabilities and utilities. Typically, there is no specific commitment concerning whether or how the relevant probabilities and utilities are represented — instead, the assumption is that preferences and subjective probabilities are "revealed" by patterns of observed choices. Indeed, given fairly natural consistency assumptions concerning how people choose, it can be shown that the observed pattern of choices can be represented in terms of expected utility — i.e., appropriate utilities and subjective probabilities can be inferred [Savage, 1954], with no commitment to their underlying psychological implementation. Indeed, this type of result can sometimes be used as reassurance that the expected utility framework is appropriate, even in complex real-world decisions, where people are unable to estimate probabilities or utilities.

The descriptive study of how people make decisions has, as with the study of reasoning, taken the normative perspective as its starting point; and aimed to test experimentally how far normative assumptions hold good. In a typical experiment, outcomes are made as clear as possible: for example, people may choose between monetary gambles, with known probabilities; or between gambles and fixed amounts of money.

A wide range of systematic departures from the norms of expected utility are observed in such experiments, as demonstrated by the remarkable research programme initiated by Kahneman, Tversky and their colleagues (e.g., [Kahneman *et al.*, 1982; Kahneman and Tversky, 2000]). Thus, for example, people can be induced to make different decisions, depending on how the problem is "framed." Thus, if a person is given £10 at the outset, and told that they must choose either a gamble, with a 50% chance of keeping the £10, and a 50% chance of losing it all; or they must give back £5 for certain, they tend to prefer to take the risk. But if they are given no initial stake, but asked whether they prefer a 50-50 chance of £10, or a certain £5, they tend to play safe. Yet, from a formal point of view these choices are identical — the only difference is that in one case the choice is framed in terms of losses (where people tend to be risk-seeking); rather than gains (where they tend to be risk-averse).

Expected utility theory cannot account for framing effects of this type — only the formal structure of the problem should matter, from a normative point of view; the way in which it is described should be irrelevant. Indeed, expected utility theory can't well account for the more basic fact that people are not risk neutral (i.e., neutral between gambles with the same expected monetrary value) for small stakes [Rabin, 2000]. This is because, from the standpoint of expected utility theory, people ought to evaluate the possible outcomes of a gamble in "global" terms — i.e., in relation to the impact on their life overall. Hence, if a person has an initial wealth of £10,000, then both the gambles above amount of choosing between a 50-50 chance of ending up with a wealth of £10,010 or £10,000, or a certain wealth of £10,005.

One reaction to this type of clash between human behaviour and rational norms is the observation that the human behaviour is error-prone — and hence, where this is true, expected utility will be inadequate as a *descriptive* theory of choice. A natural follow-up to this, though, is to attempt to modify the normative theory so that it is provides a better fit with the empirical data. A wide range of proposals of this sort have been put forward, including prospect theory [Kahneman and Tversky, 1979; Tversky and Kahneman, 1992], regret theory [Loomes and Sugden, 1982], and rank-dependent utility theory [Quiggin, 1993]. Indeed, prospect theory, by far the most influential framework, was deliberately conceived as an attempt to find "the minimal set of modifications of expected utility theory that would provide a descriptive account" of risky choices ([Kahneman, 2000, p. 411], as cited in [Brandstätter *et al.*, 2006]).

In essence, prospect theory modifies expected utility theory in three main ways. First, monetary outcomes are consider in isolation, rather than aggregated as part of total wealth. This fits with the wider observation that people tend to view different amounts of money, or indeed goals, quantities or events of any kind, one-by-one, rather than forming a view of an integrated whole. This observation is the core of Thaler's [1985] "mental accounting" theory of how people make real-world financial decisions.

Second, prospect theory assumes that while the value function (i.e., relating money to subjective value) for positive gains is concave (indicating risk aversion in an expected utility framework), the value function for losses is convex. This implies that the marginal extra pain for an additional unit of loss (e.g., each extra pound or dollar lost) decreases with the size of the loss. Thus, people are risk-seeking when a gamble is framed in terms of losses, but risk averse when it is framed in terms of gains, as we noted above. Moreover, the value function is steeper for losses than for gains, which captures the fact that most people are averse to gambles with a 1/2 chance of winning £10, and a 1/2 chance of losing -£10 [Kahneman and Tversky, 1979]. This phenomenon, *loss aversion*, has been used to explain a wide range of real world phenomena, including the status quo bias (losing one thing and gaining another tends to seem unappealing, because the loss is particularly salient, [Samuelson and Zeckhauser, 1988]) and the equity premium puzzle (share returns may be "unreasonably" high relative to fixed interest bonds, because people dislike falls in stock prices more than they like the equivalent gains, [Benartzi and Thaler, 1995]).

The final key modification of expected utility theory is that prospect theory assumes that people operate with a distorted representation of probability. They overestimate probabilities near zero; and underestimate probabilities near 1, such that the relation between probability, $p(i)$ and the "decision weights", $w(i)$, which are assumed to determined people's choices, as related by an inverse-S shape. According to prospect theory, this distortion can explain the so-called "four-fold pattern" of risky decision making — that, for small probabilities, risk-preferences reverse both for gains and losses. So for example, when probabilities are high, e.g., .5, people prefer a certain gain of £500 to the probable gain of £1000, but they prefer the probable loss of £1000 to the certain loss of £500. When probabilities are low, e.g., .0005, people prefer a probable gain of £1000 to the certain gain of 50p, but they prefer the certain loss of £500 to the probable loss of £1000.

The machinery of prospect theory integrates values and decision weights to assign a value to each gamble (where this is any choice with an uncertain outcome), just as in expected utility theory, so that the value of a risky option is:

$$\sum_i w(i)v(i)$$

where $w(i)$ is the decision weight (i.e., distorted probability) for outcome i; and $v(i)$ is the value of that outcome.

Psychological Models not Rooted in Economics

Prospect theory and other variants of expected utility, hold with the assumption that people represent value and probability on some kind of absolute internal scale; and that they integrate these values by summing the product of weight and value over possible outcomes, to obtain the value of each gamble.

Two recent psychological theories, however, set aside the structure of expected utility theory; they are inspired not by the attempt to modify normative considerations, but instead to trace the consequences of assumptions about the cognitive system.

One recent approach [Brändstatter *et al.*, 2006] focuses on processing limitations, and on the consequences of assuming that the cognitive system is not able to integrate different pieces of information, and that, instead, people can only focus on one piece of information at a time. This assumption is controversial. In perceptual judgements (e.g., concerning the identity of a phoneme, or the depth of a surface), many theories explicitly assume (linear) integration between difference sources of information [Massaro, 1987; Schrater and Kersten, 2000] — in a probabilistic framework, this corresponds, roughly, to adding logs of the strength of evidence provided by each cue. Note, moreover, that such cue integration appears to be computationally natural in neural hardware (e.g., [Deneve *et al.*, 2001]). Many models of higher-level judgement have assumed that information is also integrated, typically linearly (e.g., [Anderson, 1981; Hammond, 1996]). However, Gigerenzer and colleagues (e.g., [Gigerenzer and Goldstein, 1996; Gigerenzer *et al.*, 1999]) have influentially argued that high-level judgements — most famously, concerning the larger of pairs of German cities — do not involve integration. Instead judgement is assumed to involve considering cues, one at a time — if a cue determines which city is likely to be larger, that city is selected; if not, a further cue is chosen, and the process is repeated. There has been considerable, and ongoing, controversy concerning the circumstances under which integration does or does not occur, in the context of judgement [Hogarth and Karelaia, 2005a; 2005b].

Brändstatter, Gigerenzer and Hertwig's [2006] innovation is to show that a non-integrative model can make in-roads into understanding how people make risky decisions — a situation which has been viewed as involving the trade-off between "risk" and "return" almost by definition. Their model, the priority heuristic, has the following basic form. For gambles which contain only gains (or £0), the heuristic recommends considering features of the gambles in the order: minimum gain, probability of minimum gain, maximum gain. If gains differ by at least 1/10 of the maximum gain (or, for comparison of probabilities, if probabilities differ by at least 1/10), choose the gamble which is "best" on that feature (defined in the obvious way). Otherwise move the next feature in the list, and repeat.

To see how this works, consider the gambles used above to illustrate the "fourfold" pattern of risky choice, described by Kahneman and Tversky [1979]. For the high probability gamble over gains, the minimum gain for the certain outcome is £500; but the minimum gain for the risky gamble is £0; this difference is far more than 1/10 of the maximum gain, £1000. Hence, the safe option is preferred. By contrast, for the low probability gamble, the difference between the minimum gains for the options is just 50p, which is much less than 1/10 of he maximum gain of £1000. Hence, this reason is abandoned, and we switch to probability of minimum gain — this is clearly higher for a certain gamble — as there is only one outcome, which is by definition the minimum. The risky gamble, with the smaller probability of minimum gain, is therefore preferred. Thus, we have risk seeking with small probabilitities of large gains (and hence an explanation of why people buy lottery tickets).

Brändstatter, Gigerenzer and Hertwig propose a modification of the heuristic for gambles containing just losses, where "gain" is replaced by "loss" throughout, so that the feature order is: minimum loss, probability of minimum loss, maximum loss. If gains differ by at least 1/10 of the maximum loss (or probabilities differ by at least 1/10), choose the gamble which is "best" on that feature (defined in the obvious way). Otherwise move the next feature in the list, and repeat. Tracing through the reasoning described above, for the "loss" gambles in the "four-fold" pattern of risky choice, shows that people should appear risk seeking for losses, except where there is a small probability of a large loss; here people will again be risk averse (e.g., they will buy insurance).

The priority heuristic model does, however, make some extremely strong and counterintuitive predictions — e.g., that if the minimum gains differ sufficiently, then all other features of the gambles (including the probability of obtaining those gains) will have no impact on choice. In extreme cases, this seems implausible. For example, a certain 11p whould be preferred to a .999999 probability of £1 (and otherwise £0). Brändstatter, Gigerenzer and Hertwig [2006] restrict their account, however, to cases for which the expected values of the gambles are roughly comparable — where they are not, the gamble with the obviously higher expected value is chosen, and the priority heuristic is not invoked.

Another recent approach to risk decision making, starting from cognitive principles rather than a normative economic account, is Decision by Sampling (DbS), [Stewart *et al.*, 2006], see also [Stewart and Simpson, in press]. The starting point of DbS is the psychophysical observation that people can accurately make binary comparisons concerning the louder, or brighter, of two sensory magnitudes, but are extremely poor at judging the absolute magnitudes of such stimuli. Thus, for example, people can typically assign sensory magnitudes, however widely spaced, to no more than about five classes [Miller, 1956]; and even these crude judgements are subject to influences of the previous stimuli (e.g., [Garner, 1953]).

Indeed, to a first approximation, people's judgements can be well modeled by assuming that they have little or no coding of absolute magnitudes; but merely make relative judgements based on the "jumps" between successive magnitudes (for a detailed model along these lines, see [Stewart *et al.*, 2005]).

It seems natural to assume that representation of non-sensory magnitudes may behave similarly. If so, then the "gut feel" of how much value is associated with a particular amount of money or a particular probability may be dissociated from the absolute quantities involved. Instead, the DbS framework argues that such magnitudes are judged against a small number of other similar magnitudes, derived either from immediate context, or from memory. The rank of an item is, on this view, all that influences it subjective representation. Thus, if people have been thinking about small sums of money, a medium sized sum of money may seem large; if they have been thinking about larger sums, the same medium size sum may seem small.

This viewpoint assumes that there people have no underlying internal "scales" for utility or probability — but nonetheless, it turns out to be possible to reconstruct something analogous to the value and decision weight functions from prospect theory. If people assess the gut feel of a magnitude in relation to prior examples, the statistical distribution of such magnitudes is likely to be important. Other things being equal, this distribution will provide an estimate of the probabilities of different comparison items being considered in particular judgements. Thus, if small sums of money are much more commonly encountered than large sums of money, then it is much more likely that people will consider small sums of money as comparison items, other things being equal. Therefore, the difference in "gut" feel between £5 and £50 will be much greater than that between £1005 and £1050, because sampling an item in the first interval (so that the lower and upper items will be assigned different ranks), is much more likely than sampling in the second. More generally, the attractiveness of an option, according to DbS, is determined by its rank in the set of comparison items; and hence, its typical attractiveness (across many sampling contexts) can be estimated by its rank position in a statistical sample of occurrences of the relevant magnitude.

To examine this hypothesis, Stewart, Chater, and Brown [2006] examined a sample of "positive" sums of money — credits into accounts from a high street bank — and showed that plotting monetary value against rank produces a concave function, reminiscent of those in utility theory and prospect theory. Thus, the "gut" attractiveness of a sum of money is, on average, a diminishing function of amount. The similar analysis for losses (using bank account debits as a proxy) yields a convex function of value against losses, as in prospect theory. Moreover, for losses, the statistical distribution is more skewed towards small items, which has the consequence that ranks change more rapidly for small values for losses than for gain. This corresponds to a steeper value curve for losses and gains, and hence captures loss aversion. Indeed, putting the curves of rank against value together yields a curve strikingly reminiscent of that postulated in prospect theory.

Applying the same logic to probability requires estimating typical probabilities that people consider. Stewart *et al.*, [2006] attempt this by recording the corpus frequencies of probability-related phrases (e.g., *likely*, *slight chance*, *probable*, *extremely doubtful*, and so on); and secondly asking people to assign numerical probabilities to these phrases. This analysis yielded an estimate of the probabilities that people typically consider — and, perhaps not surprisingly, these are dense near 0 and 1. According to DbS, the gut feel of how large a probability seems depends on relative rank in this distribution — yielding an inverse S-shaped curve, as in prospect theory. Thus, DbS can capture many of the insights of prospect theory, and *explain*, rather than postulate, the relevant functional forms (e.g., concerning an analog of the inverse S-shape probability weighting function in prospect theory); but it is also able to predict strong local contextual effects, which are presumed to be determined by local sampling biases (e.g., [Stewart *et al.*, 2003). The key gap in DbS is, though, the lack of a detailed theory of how sampling occurs, in any specific decision making context.

Most decision making research has concentrated on verbally stated "one-shot" problems. But there has been a long tradition in psychology of studying how people (and animals) make repeated decisions, typically under some schedule of reinforcement [Shanks, 1995], which has led to a range of computational models, many within a Bayesian, or partially, Bayesian framework [Kruschke, 2006]. There has also been recent interest in directly comparing "decision-by-experience" with performance on descriptive decision problems [Hertwig *et al.*, 2004]. Early results appear to indicate that behaviour is different — for example, it has been argued that people may *under* rather than *over* weight small probabilities, when learning from experience (although see [Fox and Hadar, 2006]). This work is particularly interesting in the light of recent Bayesian models of reinforcement learning in animals and humans (e.g., [Courville *et al.*, 2006]; and imaging studies which are beginning to connect Bayesian decision making models with brain function [Daw *et al.*, 2006]. If aspects of human, and even rat, learning are Bayesian, the normative failure of human choice in descriptive problems seem all the more puzzling.

Moving further away from descriptive decision problems, there has been recent investigation of how decision problems framed in terms of perceptuo-motor tasks are performed (e.g., [Trommershäuser *et al.*, 2006]. This is particularly interesting, given the recent surge of interest in sophisticated Bayesian decision-theoretic models of perceptuo-motor control. The spirit of these models is that the motor system may implement (approximations to) highly elaborate probabilistic calculations, in order to reduce costs concerning energy consumption, motor error, or a costs learned direct from experienced in an experimental set-up (e.g., [Körding and Wolpert, 2006]). Trommershäuser *et al.* [2006] have show how people adjust the direction of pointing towards a "target," which provides monetary reward, but for which losses are incurred if the target is missed. It appears that the motor system rapidly adapts so that gain is maximized, in a way that is adapted to the intrinsic motor error involved in pointing. Again, the contrast of such apparently Bayesian behaviour with performance on descriptive choice problems is intriguing.

Summary: Decision Making

In both reasoning and decision making, indeed, there is a certain air of paradox in human performance [Oaksford and Chater, 1998]. Human common-sense reasoning is far more sophisticated than any current artificial intelligence models can capture; yet people's performance on, e.g., simple conditional inference, while perhaps explicable in probabilistic terms, is by no means effortless and noise-free; and similarly, in decision making, it appears that "low-level" repeated decision making may be carried out effectively (where, in the context of motor control, the complexity of the decision problem of planning trajectories for the motor system typically far exceed the capabilities of current methods [Todorov, 2004]). But perhaps this situation is not entirely paradoxical. It may be that both human reasoning and decision-making function best in the context of highly adapted cognitive processes such as basic learning, deploying world knowledge, or perceptuo-motor control. Indeed, what is striking about human cognition is the ability to handle, even to a limited extent, reasoning and decision making in novel, hypothetical, verbally stated scenarios, for which our past experience and evolutionary history may have provided us only minimal preparation.

6 ARGUMENTATION

Reasoning and decision making often takes place in the service of argumentation, i.e., the attempt to persuade yourself or others of a particular, perhaps controversial, position [van Eemeren and Grootendorst, 1992]. Argumentation is the overarching human activity that studies of deductive reasoning, inductive reasoning, judgment and decision making are really required to explain. So one might attempt to persuade someone else to accept a controversial standpoint p by trying to persuade them that p is actually a logical consequence of their prior beliefs or current commitments; or that p has strong inductive support; or, where p is an action, that p will help to achieve their, our, or the country's current goals. Recently, a Bayesian inductive logic approach has been extended to at least some aspects of argumentation (e.g., [Hahn and Oaksford, 2007]). The approach is very similar to accounts of conditional inference we reviewed above. We are concerned with the 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 of a fallacy seem to be good arguments while others do not. Fallacies — arguments that seem correct but aren't, e.g., denying the antecedent — have been a longstanding focus of debate. Catalogues of reasoning and argumentative fallacies originate with Aristotle and populate books on logic and informal reasoning to this day. The classic tool brought to the analysis of fallacies is formal logic and it is widely acknowledged to have failed in providing a satisfactory account. Testament to this is the fact that fallacies figure in logic textbooks under the header of 'informal reasoning fallacies' (see e.g., [Hamblin, 1970]) — an acknowledgement of the inability to provide a sufficient formal logical treatment. In particular, logical accounts have proved unable to capture the seeming exceptions to fallacies that arise with simple changes in content that leave the structure of the argument unaffected. This suggests that either it is not formal aspects of fallacies that make them fallacious, or else that the relevant formal aspects are not being tapped into by classical logics.

Oaksford and Hahn [2004], see also, [Hahn and Oaksford, 2006; 2007; Hahn *et al.*, 2005a; Hahn *et al.*, 2005b] provided evidence of such variation and put forward an alternative, Bayesian account: individual arguments are composed of a conclusion and premises expressing evidence for that conclusion. Both conclusion and premises have associated probabilities which are viewed as expressions of subjective degrees of belief. Bayes' theorem then provides an update rule for the degree of belief associated with the conclusion in light of the evidence. *Inductive strength*, then, on this account is a function of the degree of prior conviction, the probability of evidence, and the relationship between the claim and the evidence, in particular how much more likely the evidence would be if the claim were true. That is, different instances of argumentative fallacies may vary in inductive strength conceived of as the probability of the conclusion given the premises. Oaksford and Hahn [2007] also show how the concept of inductive strength in argumentation is related to the probabilistic analysis of the conditional (see above) and recent discussion in Rips (2001). We illustrate this approach by appeal to a particular informal reasoning fallacy: the argument from ignorance.

A Probabilistic Approach

A classic informal argument fallacy, which dates back to John Locke, is the so-called argument from ignorance, or *argumentum ad ignorantiam.*

(7) Ghosts exist, because nobody has proven that they don't.

This argument does indeed seem weak. One would hesitate in positing the existence of all manner of things whose non-existence simply had not been proven, whether these be UFO's or flying pigs with purple stripes. However, is it really the general structure of this argument that makes it weak, and if so what aspect of it is responsible? Other arguments from negative evidence are routine in scientific and everyday discourse and seem perfectly acceptable:

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

Should all arguments from negative evidence be avoided, or can a systematic difference between the two examples be recognized and explained?

A Bayesian account can capture the difference between (7) and (8) as we show below. Moreover, it can capture the difference between positive and negative evidence which allows one to capture the intuition that the positive argument (9) is stronger than the negative argument (10):[3]

(9) Drug A is toxic because a toxic effect was observed (positive argument)

(10) Drug A is not toxic because no toxic effects were observed (negative argument, i.e., the argument from ignorance).

Though (10) too can be acceptable where a legitimate test has been performed, i.e.,

[3]One might argue that (9) and (10) are problematic because replacing "not toxic" with "safe" would alter the status of these arguments. This is not the case because we do not have a concept of a "safe effect." The *tests* are tests for *toxic* effects. So (10) could be rephrased as, "Drug A is safe because no toxic effects were observed," but not as, "Drug A is safe because *safe* effects were observed." As the observation of toxic effects is driving these distinctions, what "safe" means in this context must be defined in terms of toxicity in order to define the relevant probabilities.

If drug A were toxic, it would produce toxic effects in legitimate tests.
Drug A has not produced toxic effects in such tests
Therefore, A is not toxic

Demonstrating the relevance of Bayesian inference for negative vs. positive arguments involves defining the conditions for a legitimate test. Let e stand for an experiment where a toxic effect is observed and $\neg e$ stands for an experiment where a toxic effect is not observed; likewise let T stand for the hypothesis that the drug produces a toxic effect and $\neg T$ stand for the alternative hypothesis that the drug does not produce toxic effects. The strength of the argument from ignorance is given by the conditional probability that the hypothesis, T, is false given that a negative test result, $\neg e$, is found, $P(\neg T|\neg e)$. This probability is referred to as negative test validity. The strength of the argument we wish to compare with the argument from ignorance is given by positive test validity, i.e., the probability that the hypothesis, T, is true given that a positive test result, e, is found, $P(T|e)$. 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, i.e., $n = P(e|T)$, l denote selectivity, i.e., $l = P(\neg e|\neg T)$, and h denote the prior probability of drug A being toxic, i.e., $h = P(T)$, then,

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

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

Sensitivity corresponds to the "hit rate" of the test and 1 minus the selectivity corresponds to the "false positive rate."

Positive test validity is greater than negative test validity as long as the following inequality holds:

(13) $$h^2(n - n^2) > (1-h)^2(l - l^2)$$

Assuming maximal uncertainty about the toxicity of drug A, i.e., $P(T) = .5 = h$, this means that positive test validity, $P(T|e)$, is greater than negative test validity, $P(\neg T|\neg e)$, when selectivity (l) is higher than sensitivity. As Oaksford and Hahn [2004] argue, this is often a condition met in practice for a variety of clinical and psychological tests. Therefore, in a variety of settings, positive arguments are stronger than negative arguments.

The Empirical Data

Oaksford and Hahn [2004] provided experimental evidence to the effect that positive arguments such as (9) are indeed viewed as more convincing than their negative counterparts under the conditions just described. The evidence from their experiment further showed that people are sensitive to manipulations in the amount of evidence (one versus 50 studies or tests) as predicted by the account. Finally, participants' in their experiment displayed sensitivity to the degree of prior belief a character in a dialogue initially displayed toward the conclusion as the Bayesian account predicts. This finding captures the 'audience dependence' of argumentation assumed in the rhetorical research tradition (e.g., [Perelman and Olbrechts-Tyteca, 1969]).

Hahn *et al.* [2005a] generalised this account to other versions of the argument from ignorance and addressed an outstanding problem. The ghosts example (14) differs from Oaksford and Hahn's [2004] experimental materials in one, possibly important way. The argument for ghosts not only involves negative evidence, but also a flip in polarity between evidence and conclusion: negative evidence is provided to support the *positive* existence of something. In other words the inference is of the form:

(14) *not* proven (*not* exist) → exist

as opposed to merely:

(15) *not* proven (exist) → *not* exist

The examples in Oaksford and Hahn [2004] have the structure in (15) not the structure in (14). But it may be the opposite polarity case (14) that constitutes the true fallacy of the argument from ignorance.

Classical logic licenses an inference from *not*(*not* p) to p, but not the inference underlying (14) which might be rendered as:

(16) *not* says (*not* p) →?

This is because when one has not said '*not* p,' one can either have said 'p' or not spoken about 'p' at all. For example, in an argument one might defend oneself with the claim "I didn't say you were rude", which could be true either because one had specifically claimed the opposite or because one had not mentioned rudeness at all. So maybe nothing at all can be inferred in such cases?

Hahn *et al.* [2005a] established that (16) can be a strong argument by using a form of the argument from ignorance based on *epistemic closure* which is related to the negation as failure procedure in artificial intelligence [Clark, 1978]. The case can be made with an informal example: imagine your work colleagues are having a staff picnic. You ask the picnic organizer whether your colleague Smith is coming and receive the reply that "Smith hasn't said that he's not coming". Should this allow you to infer that he is in fact coming, or has he simply failed to send the required reply by e-mail? Your confidence that Smith will be attending will vary depending on the number of people that have replied. If you are told that no one has replied so far, assuming Smith's attendance seems premature; if by contrast you are told that everyone has replied, you would be assured of his presence. In between these two extremes your degree of confidence will be scaled: the more people have replied the more confident you will be. In other words, the epistemic closure of the database in question (the e-mail inbox of the organizer) can vary from no closure whatsoever to complete closure, giving rise to corresponding changes in the probability that *not says* (*not p*) does in fact suggest that *p*.

Hahn *et al.* [2005a] experiments confirmed that people are sensitive to variations in the epistemic closure of a database and that this affects their willingness to endorse argument like (16). Moreover, they found that arguments like (16) can be regarded as stronger than the standard negative evidence case (15). Therefore, as our example suggested, there would seem to be nothing in the structure of arguments like the Ghosts example that make them inherently unacceptable.

The real reasons why negative evidence on ghosts is weak, i.e., why (7) is a weaker argument than (8), are the lack of sensitivity (ability to detect ghosts) of our tests as well as our low prior belief in their existence, i.e., (7) is weak a because of the probabilistic factors that affect the strength of the argument. Hahn and Oaksford [2006; 2007] have shown how this account generalises to other inferential fallacies, such as *circularity* and the *slippery slope* argument.

Summary: Argumentation

In summary, in this section we have shown how comparing arguments in terms of their inductive strength can resolve the problem of why some instances of informal argument fallacies nonetheless seem like perfectly acceptable arguments that should rationally persuade an audience.

7 CHALLENGES AND FUTURE DIRECTIONS

The models we have discussed in this review generally treat probabilistic methods as shedding important light on cognitive processes, although in a variety of ways, and at a variety of levels of explanation, as we have seen. Yet these applications of probability can, individually and collectively, be criticized — and the debates between proponents of probabilistic methods, and advocates of alternative viewpoints, have played an important role in the development of the cognitive sciences; and are likely to continue to do so. We briefly here consider some of the many concerns that may be raised against probabilistic approaches.

Probabilistic approaches may be especially vulnerable, as noted above, when considered as models of explicit reasoning. As we have mentioned, there have been repeated demonstrations that explicit human decision making systematically deviates from Bayesian decision theory [Kahneman *et al.*, 1982; Kahneman and Tversky, 2000]. Why might such deviations occur? Since Simon [1957], computational tractability has been a primary concern — with the conclusion that computationally cheap heuristic methods, which function reasoning well in the ecological environment in which the task must be performed, should be viewed as an alternative paradigm. Bounded rationality considerations have gradually become increasingly important in economics (e.g., [Rubinstein, 1998] — and hence, economists have increasingly begun to question the usefulness of strong rationality assumptions, such that agents are viewed as implicit probabilists and decision theorists. Gigerenzer [Gigerenzer *et al.*, 1999]) has led a particularly influential programme of research, aiming to define an “ecological” rationality, in which good reasoning is that which works quickly and effectively in the real world, rather than necessarily being justified in terms of normative mathematical foundations. This viewpoint may still see a role for probabilistic analysis — but as providing an explanation of why particular heuristics work in particular environments, rather than as characterizing the calculations that the cognitive system performs (a similar approach is adopted in the probability heuristics model of quantified syllogistic reasoning).

A very different reason why people may not, in some contexts, be viewed as probabilists or decision theorists, concerns *representation*, rather than processing power. Some researchers (e.g., [Laming, 1997]) argue that people can only represent sensory magnitudes in relative rather than absolute terms; and that even this relative coding is extremely inaccurate and unstable. Indeed, the radical assumption that, to an approximation, people can make only simple qualitative binary judgements (e.g., “tone A is louder than tone B”; and “the difference in loudness between tones A and B is smaller than the difference in loudness between tones B and C”) is the basis for a recent model, the Relative Judgment Model [Stewart *et al.*, 2005] that provides a simple and comprehensive account of how people can assign sensory magnitudes to discrete categories. If the same principles apply to magnitudes involved in decision making (e.g., time, probability, value, quality, and so on), then this suggests that people may not have a stable cardinal representation of the relevant decision variables, from which probabilistic calculations (of expected utility and the like) must begin — and hence the issue of computational considerations does not even arise. The recent model of risky decision making, Decision by Sampling [Stewart *et al.*, 2006] mentioned above, shows how the assumption that people have no access to internal scales, but rely instead purely in binary judgments, can provide a straightforward account many well-known phenomena concerning risky decision making. This type of approach is extended to consider how far anomalies of choice in which items have multiple dimensions, which must be traded off, can explained in this framework.

The concern that people do not have the appropriate representations over which probabilistic calculations can be performed may be most pressing in the context of explicit reasoning — where the underlying computational machinery has not been finely adapted over a long evolutionary history to solve a stable class of problems (e.g., such as perceiving depth, or reaching and grasping) but rather the cognitive system is finding an *ad hoc* solution, as best it can, to each fresh problem. Thus, as noted above, we may accept that explicit reasoning with probability may be poor, while proposing that underlying computational processes of perception, motor control, learning and so on, should be understood in probabilistic terms.

Interestingly, though, related challenges to the Bayesian approach have arisen in perception. For example, Purves and colleagues (e.g., [Howe and Purves, 2004; 2005; Long and Purves, 2002; Nundy and Purves, 2002]) argue that the perceptual system should not be viewed as attempting to reconstruct the external world using

Bayesian methods. Instead, they suggest that the output of the perceptual system should be seen as determined by the ranking of the present input in relation to the statistical distribution of previous inputs. This viewpoint is particularly clearly expressed in the context of lightness perception. The perceived lightness of a patch in the sensory array is determined not merely by the amount of incident energy in that patch, and its spectral composition, but is also a complex function of the properties of the area surrounding that patch. For example, a patch on the sensory array will be perceived as light if it is surrounded by a dark field; and may be perceived as relatively dark, if surrounded by light field.

A natural Bayesian interpretation of this type of phenomena is that the perceptual system is attempting to factor out the properties of the light source, and to represent only the reflectance function of the surface of the patch (i.e., the degree to which that patch absorbs incident light). Thus, the dark surrounding field is viewed as prima facie evidence that the lighting is dim; and hence the patch itself is viewed as reflective; a bright surrounding field appears to support the opposite inference. This type of analysis can be formulated elegantly in probabilistic terms [Knill and Richards, 1996]. Purves and colleagues argue, instead, that the percept should not be viewed as reconstructing an underlying reflectance function — or indeed any other underlying feature of the external world. Instead, they suggest that the background field provides a context in which statistics concerning the amount of incident light is collected; and the lightness of a particular patch, in that context, is determined by its rank in that statistical distribution. Thus, when the surround is dark, patches within that surround tend to be dark (e.g., because both may be explained by the presence of a dim light source); when the surround is light, patches in that surround tend to be light. Hence, the rank position of an identical patch will differ in the two cases, hence leading to contrasting lightness percepts. Nundy and Purves [2002] conduct extensive analysis of the statistical properties of natural images, and argue that the resulting predictions frequently depart from the predictions of the Bayesian analysis; and that the rank-based statistical analysis better fits the psychophysical data.

Various responses from a Bayesian standpoint are possible — including, most naturally, the argument that, where statistical properties of images diverge from the properties of an underlying probabilistic model, this is simply an indication that the probabilistic model is incomplete. Thus, a revised Bayesian approach may account for apparent anomalies, as the model should more accurate capture the statistical properties of images. To some degree, this response may seem unsatisfying, as the ability to choose between the enormous variety of probabilistic image models may seem to give the Bayesian excessive theoretical latitude. On the other hand, the choice of model is actually strong constrained, precisely because its output can directly be tested, to see how far it reproduces the statistical properties of natural images [Yuille and Kersten, 2006]. But the challenge of the Purves's approach is that the probabilistic machinery of the Bayesian approach is unnecessary — that there is a much more direct explanation of perceptual experience, which does not involve factoring apart luminance levels and reflectance functions; but which works directly with the amount of incident light in field and surround; and which considers only ordinal properties of relevant statistical distributions, rather than the absolute magnitudes that appear to be the appropriate to a Bayesian analysis. Whether such calculations should best be viewed as departing entirely from the probabilistic approach, or rather as an illustration of how probabilistic calculations can be approximated cheaply, by analogy with heuristic-based approaches to decision making, is not clear.

A more general objection to the probabilistic approach to cognition, which we have touched on already, is the complexity of the approach. In one sense, the probabilistic approach is elegantly simple — we need simple assign prior probabilities, and then remorsely follow the laws of the probability calculus, as further data arises. But in another sense, it is often highly complex — because assigning priors to patterns of belief, images, or sentences, may require specifying an extremely complex probabilistic model, from which such information can be generated. Thus, the cognitive modeller may sometimes be accused of putting so much complexity into the model that the ability to capture the relevant data is hardly impressive. This chapter illustrates that the balance between model and data complexity is not necessarily out of balance. Moreover, the contribution of Bayesian models may often be in providing qualitative explanations (e.g., for why there should be a direct relationship between the probability of recurrence of an item, and its retrievability from memory, e.g., [Anderson and Milson, 1989; Anderson and Schooler, 1991, Schooler and Anderson, 1997].

Despite this, however, the question of how to constrain probabilistic models as far as possible in an important one. One approach, for example, is to take representation, rather than probability, as the basic construct. According to this approach, the preferred interpretation of a set of data is that which can be used to provide the shortest encoding of that data. Thus, the problem of probabilistic inference is replaced by a problem of finding short codes. It turns out that there are very close relationships between the two approaches, based in both Shannon's theory of communication [Shannon and Weaver, 1949; Mackay, 2003]; and the more general concept of algorithmic information, quantified by Kolmogorov complexity theory [Li and Vitányi, 1997]. These relationships are used to argue that the two approaches make identical behavioral predictions [Chater, 1996]. Roughly, the idea is that representations may be viewed as defining priors, such that, for any object x, with a shortest code of length $c(x)$, the prior $\Pr(x)$ is $2^{-c(x)}$. Conversely, for any prior distribution $Q(x)$ (subject to mild computability constraints that need not detain us here), there will be a corresponding system of representation (i.e., a coding language) c_Q, such that, for any data, x, probable representations or hypotheses, H_i, will correspond to those which provide the shortest codes for x. This means, roughly, that the probabilistic objective of finding the most probable hypothesis can be replaced by the coding objective of finding the hypothesis that supports the shortest code. The equivalence of these frameworks can be viewed as resolving a long-standing dispute between simplicity and likelihood (i.e., probabilistic) views of perceptual organization (e.g., [Pomerantz and Kubovy, 1987]), as argued by Chater [1996].

Despite these close relationships, taking representation and coding as basic notions has certain advantages. First, the cognitive sciences arguable already have a substantial body of information concerning how different types of information is represented — certainly this has been a central topic of experimental and theoretical concern; but by contrast the project of assessing probabilistic models directly seems more difficult. Second, priors are frequently required for representations which presumably have not been considered by the cognitive system. In a standard Bayesian framework, we typically define a space of hypotheses, and assign priors over that space; but we may also wonder what prior would be assigned to a new hypothesis, if it were considered (e.g., if a particular pattern is noticed by the perceptual system; or if a new hypothesis is proposed by the scientist). Assuming that the coding language is universal, then these priors are well-defined, even for an agent that has not considered them — their prior probability of any H is presumed to be $2^{-c(H)}$. Third, rooting priors in a coding language frees the cognitive system from the problem of explicitly having to represent such prior information (though this may be done in a very elegant and compact form, see, e.g., [Tenenbaum *et al.*, 2006]).

Technical developments in coding-based approaches to inference (e.g., [Barron *et al.*, 1998; Hutter, 2004; Li and Vitányi, 1997; Rissanen, 1987; 1996; Wallace and Freeman, 1987]) as well as applications to cognition (e.g., [Brent and Cartwright, 1996; Chater and Vitányi, 2007; Dowman, 2000; Feldman, 2000; Goldsmith, 2001; Pothos and Chater, 2002]) have been divided concerning whether a coding-based approach to inference should be viewed as a variant of the probabilistic account (i.e., roughly, as using code lengths as a particular way of assigning priors); or whether it should be viewed as an alternative approach. One argument for the former, harmonious, interpretation is that the probabilistic interpretation appears necessary if we consider choice. Thus, for example, maximizing expected utility (or similar) requires computing expectations — i.e., knowing the probability of various outcomes. Thus, rather than viewing simplicity-based approaches as a rival to the probabilistic account of the mind, we instead tentatively conclude that it should be viewed as an alternative, and often useful, perspective on probabilistic inference.

CONCLUSION

This chapter has introduced the ways in which inductive logic has been applied in empirical psychology to provide models of a range of high level cognitive abilities. Although Bayesian methods have been applied at a variety of levels of explanation of cognition and perception, we have concentrated in the main on central cognitive processes [Fodor, 1983]. These are the processes of central concern in philosophical logic, i.e., those where the inferences involved can be expressed verbally and where a clear delineation between premises and conclusion can be made. In language, inductive reasoning, deductive reasoning, argumentation, and decision making, we have shown that inductive logic has been able to provide new insights in to the processes involved. Thus, in recent years it seems inductive logic has facilitated many promising developments in the attempt to understand human cognition.

ACKNOWLEDGEMENTS

Acknowledgements: Nick Chater is supported by a Senior Research Fellowship from the Leverhulme Trust, and the ESRC Centre for Economic Learning and Social Evolution (ELSE).

BIBLIOGRAPHY

[Adams, 1975] E. Adams. *The logic of conditionals: An application of probability to deductive logic*. Dordrecht: Reidel, 1975.

[Adams, 1998] E. Adams. *A primer of probability logic*. Stanford, CA: CSLI Publications, 1998.

[Adelson, 1993] E. H. Adelson. Perceptual organization and the judgment of brightness. *Science*, 262, 2042-2044, 1993.

[Adelson and Pentland, 1996] E. H. Adelson and A. P. Pentland. The perception of shading and reflectance. In D. Knill and W. Richards (Eds.) *Perception as Bayesian Inference*. Cambridge University Press, pp. 409-423, 1996.

[Akaike, 1974] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723, 1974.

[Anderson, 1990] J. R. Anderson. *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1990.

[Anderson, 1991a] J. R. Anderson. Is human cognition adaptive? *Behavioral and Brain Sciences, 14*, 471-517, 1991.

[Anderson, 1991b] J. R. Anderson. The adaptive nature of human categorization. *Psychological Review, 98,* 409-429, 1991.

[Anderson and Matessa, 1998] J. R. Anderson and M. Matessa. The rational analysis of categorization and the ACT-R architecture. In M. Oaksford and N. Chater (Eds.), *Rational models of cognition* (pp.197-217). Oxford, England: Oxford University Press, 1998.

[Anderson and Milson, 1989] J. R. Anderson and R. Milson. Human memory: An adaptive perspective. *Psychological Review, 96*, 703–719, 1989.

[Anderson and Schooler, 1991] J. R. Anderson and L. J. Schooler. Reflections of the environment in memory. *Psychological Science*, 1, 396–408, 1991.

[Anderson, 1981] N. H. Anderson. *Foundations of information integration theory*. New York, NY: Academic Press, 1981.

[Aristotle, 1908] Aristotle. *Nicomachean Ethics* (W. D. Ross, trans.). Oxford, England: Clarendon Press, 1908.

[Attneave, 1954] F. Attneave. Some informational aspects of visual perception. *Psychological Review*,61, 183-193, 1954.

[Baayen and Moscoso del Prado Martín, 2005] R. H. Baayen and F. Moscoso del Prado Martín. Semantic density and past-tense formation in three Germanic languages. *Language* 81(3), 666-698, 2005.

[Barlow, 1959] H. B. Barlow. Possible principles underlying the transformation of sensory messages. In. W. Rosenblith (Ed.) *Sensory Communication* (pp. 217-234). Cambridge, MA: MIT Press, 1959.

[Baron, 1981] J. Baron. *An analysis of confirmation bias*. Paper presented at 22^{nd} Annual Meeting of the Psychonomic Society, 1981.

[Baron, 1985] J. Baron. *Rationality and intelligence*. Cambridge, England: Cambridge University Press, 1985.

[Barron *et al.*, 1998] A. R. Barron, J. Rissanen, and B. Yu. The minimum description length principle in coding and modeling. *IEEE Transactions on Information Theory*, *IT-44*, 2743-2760, 1998.

[Barsalou, 1987] L. W. Barsalou. The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Emory Symposia in Cognition 1, Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorisation* (pp. 101-140). Cambridge, England: Cambridge University Press, 1987.

[Barwise and Cooper, 1981] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy, 4*,159–219, 1981.

[Becker, 1976] G. Becker. *The economic approach to human behavior*. Chicago: Chicago University Press, 1976.

[Becker, 1996] G. Becker. *Accounting for tastes*. Cambridge, MA: Harvard University Press, 1996.

[Benartzi and Thaler, 1995] S. Benartzi and R. H. Thaler. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, 110, 73-92, 1995.

[Bennett, 2003] J. Bennett. *A philosophical guide to conditionals*. Oxford, England: Oxford University Press, 2003.

[Berger, 1985] J. Berger. *Statistical decision theory and Bayesian analysis*. New York, NY: Springer–Verlag, 1985.

[Bernado and Smith, 1994] J. M. Bernado and A. F. Smith. *Bayesian theory*. New York, NY: Wiley, 1994.

[Bernoulli, 1713] J. Bernoulli. *Ars conjectandi, The art of conjecturin,* (trans. and notes by E. D. Sylla). Baltimore, MD, 1713. John Hopkins University Press (2005).

[Blake *et al.*, 1996] A. Blake, H. H. Bulthoff, and D. Sheinberg. Shape from texture: ideal observers and human psychophysics. In D. Knill and W. Richards, (eds.) *Perception as Bayesian Inference* (pp. 287–321). Cambridge: Cambridge University Press, 1996.

[Blakemore, 1990] C. Blakemore. *Vision Coding and Efficiency*. Cambridge: Cambridge University Press, 1990.

[Blei *et al.*, 2004] D. M. Blei, T. L. Griffiths, M. I. Jordan, and J. B. Tenenbaum. Hierarchical topic models and the nested Chinese restaurant process. *Advances in Neural Information Processing Systems 16*, Cambridge, MA: MIT Press, 2004.

[Bod *et al.*, 2003] R. Bod, J. Hay, and S. Jannedy, eds. *Probabilistic Linguistics*, MIT Press, 2003.

[Bogacz, 2007] R. Bogacz. Optimal decision-making theories: linking neurobiology with behaviour. *Trends in Cognitive Sciences*, 11, 118-125, 2007.

[Boole, 1854] G. Boole. *An investigation of the laws of though*t. London: Macmillan, 1854. Reprinted by Dover Publications, New York (1958).

[Boolos and Jeffrey, 1980] G. Boolos and R. C. Jeffrey. *Computability and logic* (2^{nd}Edition). Cambridge, England: Cambridge University Press, 1980.

[Bovens and Hartmann, 2003] L. Bovens and S. Hartmann. *Bayesian Epistemology*. Oxford: Clarendon Press, 2003.

[Braine, 1978] M. D. S. Braine. On the relation between the natural logic of reasoning and standard logic. *Psychological Review, 85*, 1-21, 1978.

[Brandstätter *et al.*, 2006] E. Brandstätter, G. Gigerenzer, and R. Hertwig. The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113, 409-432, 2006.

[Brent and Cartwright, 1996a] M. R. Brent and T. A. Cartwright. Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61, 93-126, 1996.

[Brent and Cartwright, 1996b] M. R. Brent and T. A. Cartwright. Distributional Regularity and phonotactic constraints are useful for segmentation. *Cognition* 61:93-125, 1996.

[Brunswik, 1955] E. Brunswik. Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193–217, 1955.

[Carnap, 1950] R. Carnap. *Logical foundations of probability.* 2^{nd} Edition. Chicago: University of Chicago Press, 1950.

[Charniak, 1997] E. Charniak. Statistical parsing with a context-free grammar and word statistics. In Proceedings of the 14th National Conference on Artificial Intelligence. AAAI Press, Cambridge, MA, pages 598-603, 1997.

[Chater, 1996] N. Chater. Reconciling Simplicity and Likelihood Principles in Perceptual Organization, *Psychological Review*,103, 566–581, 1996.

[Chater, 2004] N. Chater. What can be learned from positive data? Insights from an 'ideal learner.' *Journal of Child Language*, 31, 915-918, 2004.

[Chater and Manning, 2006a] N. Chater and C. Manning. Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, 10, 287-291, 2006.

[Chater and Manning, 2006b] N. Chater and C. Manning. Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, 10, 335-344, 2006.

[Chater and Oaksford, 1999] N. Chater and M. Oaksford. The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38, 191-258, 1999.

[Chater and Oaksford, 2008] N. Chater and M. Oaksford, eds. *The probabilistic mind*: *Prospects for Bayesian cognitive science*, Oxford, England: Oxford University Press, 2008.

[Chater and Vitányi, 2007] N. Chater and P. Vitányi. 'Ideal learning' of natural language: Positive results about learning from positive evidence. *Journal of Mathematical Psychology*, 51, 135-162, 2007.

[Chater *et al.*, 1998a] N. Chater, M. Crocker, and M. Pickering. The rational analysis of inquiry: The case of parsing: In M. Oaksfield and N. Chater (Eds.), *Rational models of cognition* (pp.441-468). Oxford, England: Oxford University Press, 1998.

[Chater *et al.*, 1998b] N. Chater, M. Crocker, and M. Pickering. The rational analysis of inquiry: The case of parsing. In M. Oaksford, and N. Chater (Eds.) *Rational models of cognition* (pp. 441-469). Oxford: Oxford University Press, 1998.

[Chater *et al.*, 2006] N. Chater, J. B. Tenenbaum, and A. Yuille. Special Issue on Probabilistic Models of Cognition, *Trends in Cognitive Sciences*, 10, 287-344, 2006.

[Cheng, 1997] P. W. Cheng. From covariation to causation: A causal power theory. *Psychological Review*, 104, 367–405, 1997.

[Cheng and Holyoak, 1985] P. W. Cheng and K. J. Holyoak. Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416, 1985.

[Chomsky, 1957] N. Chomsky. *Syntactic Structures*. The Hague: Mouton, 1957.

[Chomsky, 1965] N. Chomsky. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press, 1965.

[Chomsky, 1981] N. Chomsky. *Lectures on Government and Binding*, Dordrecht: Foris, 1981.

[Christiani and Shawe-Taylor, 2000] N. Christiani and J. Shawe-Taylor. *An Introduction to Support Vector Machines*. Cambridge: Cambridge University Press, 2000.

[Christiansen and Chater, 2001] M. H. Christiansen and N. Chater, eds. *Connectionist psycholinguistics*. Westport, CT: Ablex, 2001.

[Clark, 1978] K. L. Clark. Negation as failure. In H. Gallaire and J. Minker (Eds.), *Logic and databases* (pp. 293-322). New York: Plenum Press, 1978.

[Cohen, 1981] L. J. Cohen. Can Human Irrationality Be Experimentally Demonstrated? *Behavioral and Brain Sciences, 4*, 317-370, 1981.

[Collins, 2003] M. Collins. Head-Driven Statistical Models for Natural Language Parsing. *Computational Linguistics* 29(4): 589-637, 2003.

[Copeland, 2006] D. Copeland. Theories of categorical reasoning and extended syllogisms. *Thinking and Reasoning*, 12, 379-412, 2006.

[Copeland and Radvansky, 2004] D. Copeland and G. A. Radvansky. Working memory and syllogistic reasoning. *Quarterly Journal of Experimental Psychology*, 57*A*, 1437-1457, 2004.

[Cosmides, 1989] L. Cosmides. The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31,* 187-276, 1989.

[Cosmides and Tooby, 2000] L. Cosmides and J. Tooby. Evolutionary psychology and the emotions. In M. Lewis and J. M. Haviland-Jones (Eds.), *Handbook of Emotions, 2nd Edition*. (pp. 91-115). New York, NY: Guilford, 2000.

[Courville *et al.*, 2006] A. C. Courville, N. D. Daw, and D. S. Touretzky. Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, 10, 294-300, 2006.

[Crocker, 2000] M. W. Crocker and T. Brants. Wide-coverage probabilistic sentence processing. *Journal of Psycholinguistic Research* 29, 647-669, 2000.

[Culicover, 1999] P. W. Culicover. *Syntactic Nuts*. Oxford University Press, 1999.

[Cummins, 1995] D. D. Cummins. Naïve theories and causal deduction. *Memory and Cognition, 23,* 646-658, 1995.

[Daelmans and van den Bosch, 2005] W. Daelemans and A. van den Bosch. *Memory-based language processing*. Cambridge: Cambridge University Press, 2005.

[Daston, 1988] L. Daston. *Classical probability in the enlightenment*. Princeton, NJ: Princeton University Press, 1988.

[Davidson, 1984] D. Davidson. *Inquiries into truth and interpretation*. Oxford: Oxford University Press, 1984.

[Daw *et al.*, 2006] N. D. Daw, J. P. O'Doherty, B. Seymour, P. Dayan, and R. J. Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441, 876-879, 2006.

[Dayan and Abbott, 2001] P. Dayan and L. F. Abbott. *Theoretical neuroscience: computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press, 2001.

[Deneve *et al.*, 2001] S. Deneve, P. E. Latham, and A. Pouget. Efficient computation and cue integration with noisy population codes, *Nature Neuroscience*, 4, 826-831, 2001.

[Dennis, 2005] S. Dennis. A memory-based theory of verbal cognition. *Cognitive Science*, 29, 145-193. 2005.

[Desmet *et al.*, 2006] T. Desmet, De Baecke, Drieghe, Brysbaert and Vonk. Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, 21, 453-485, 2006.

[Desmet and Gibson, 2003] T. Desmet and E. Gibson. Disambiguation Preferences and Corpus Frequencies in Noun Phrase Conjunction. *Journal of Memory and Language, 49,* 353-374, 2003.

[Dickstein, 1978] L. S. Dickstein. The effect of figure on syllogistic reasoning. *Memory and Cognition, 6,* 76-83, 1978.

[Dowman, 2000] M. Dowman. Addressing the learnability of verb subcategorizations with Bayesian inference. In L. R. Gleitman and A. K. Joshi (Eds.). *Proceedings of the Twenty Second Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Erlbaum, 2000.

[Earman, 1992] J. Earman. *Bayes or bust*? Cambridge, MA: MIT Press, 1992.

[Edgington, 1995] D. Edgington. On conditionals. *Mind*, 104, 235–329, 1995.

[Edwards, 1954] W. Edwards. The theory of decision making. *Psychological Bulletin, 41*, 380-417, 1954.

[Eemeren and Grootendorst, 1992] F. H. van Eemeren and R. Grootendorst. *Argumentation, communication, and fallacies*. Hillsdale, NJ: Lawrence Erlbaum, 1992.

[Elman, 1990] J. L. Elman. Finding structure in time. *Cognitive Science*, 14, 179-211, 1990.

[Elster, 1986] J. Elster, ed. *Rational choice*. Oxford: Basil Blackwell, 1986.

[Evans, 1972] J. St. B. T. Evans. Reasoning with negatives. *British Journal of Psychology, 63,* 213-219, 1972.

[Evans *et al.*, 2003] J. St. B. T. Evans, S. H. Handley, and D. E. Over. Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 321-355, 2003.

[Evans *et al.*, 1993] J. St. B. T. Evans, S. E. Newstead, R. J. Byrne. *Human Reasoning*, Lawrence Erlbaum Associates, Hillsdale, N.J, 1993.

[Evans and Handley, 1999] J. St. B. T. Evans and S. J. Handley. The role of negation in conditional inference. *Quarterly Journal of Experimental Psychology, 52A,* 739-769, 1999.

[Evans and Over, 1996a] J. St. B. T. Evans and D. E. Over. *Rationality and reasoning*. Psychology Press: Hove, Sussex, 1996.

[Evans and Over, 1996b] J. St. B. T. Evans and D. E. Over. Rationality in the selection task: Epistemic utility versus uncertainty reduction. *Psychological Review*, 103, 356-363, 1996.

[Evans *et al.*, 1999] J. St. B. T. Evans, S. J. Handley, C. N. J. Harper, and P. N. Johnson-Laird. Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1495-1513, 1999.

[Evans *et al.*, 2003] J. St. B. T. Evans, S. H. Handley, and D. E. Over. Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 321-355, 2003.

[Evans and Over, 2004] J. St. B. T. Evans and D. E. Over. *If.* Oxford, England: Oxford University Press, 2004.

[Fanselow *et al.*, 2006] G. Fanselow, C. Féry, R. Vogel, and M. Schlesewsky, eds. *Gradience in Grammar: Generative Perspectives.* Oxford: Oxford University Press, 2006.

[Feeney and Handley, 2000] A. Feeney and S. J. Handley. The suppression of q card selections: Evidence for deductive inference in Wason's selection task. *Quarterly Journal of Experimental Psychology, 53,* 1224-1242, 2000.

[Feldman and Singh, 2005] J. Feldman and M. Singh. Information along curves and closed contours. *Psychological Review*, 112, 243-252, 2005.

[Feldman, 2000] J. Feldman. Minimization of Boolean complexity in human concept learning. *Nature*, 407, 630–633, 2000.

[Feldman, 2001] J. Feldman. Bayesian contour integration. *Perception and Psychophysics*, 63, 1171-1182, 2001.

[Fiedler and Freytag, 2004] K. Fiedler and P. Freytag. Pseudocontingencies. *Journal of Personality and Social Psychology*, 87, 453-467, 2004.

[Fiedler and Juslin, 2006] K. Fiedler and P. Juslin. *Information sampling and adaptive cognition*. New York: Cambridge University Press, 2006.

[Fitelson, 2005] B. Fitelson. Inductive logic. In J. Pfeifer, and S. Sarkar (Eds.), *The philosophy of science*. Oxford, UK: Routledge, 2005.

[Fodor, 1983] J. A. Fodor. *Modularity of mind*. Cambridge, MA: MIT Press, 1983.

[Fodor, 1987] J. A. Fodor. *Psychosemantics*. Cambridge, MA: MIT Press, 1987.

[Fodor *et al.*, 1974] J. A. Fodor, T. G. Bever, and M. F. Garrett. *The Psychology of Language*. New York:. McGraw-Hill, 1974.

[Fox and Hadar, 2006] C. R. Fox and L. Hadar. "Decisions from experience" = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber and Erev (2004). *Judgment and Decision Making*, 1, 2006.

[Frazier and Fodor, 1978] L. Frazier and J. D. Fodor. The sausage machine: A new two-stage parsing model. *Cognition*, 13, 187-222, 1978.

[Frazier, 1979] L. Frazier. On Comprehending Sentences: Syntactic Parsing Strategies. Ph.D. Dissertation, University of Connecticut, 1979.

[Freeman, 1994] W. T. Freeman. The generic viewpoint assumption in a framework for visual perception, *Nature*, 368, 542–545, 1994.

[Garner, 1953] W. R. Garner. An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology, 46,* 373–380, 1953.

[Geman and Geman, 1984] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,6, 721-741, 1984.

[Geurts, 2003] B. Geurts. Reasoning with quantifiers. *Cognition, 86*, 223-251, 2003.

[Gibson and Wexler, 1994] E. Gibson and K. Wexler. Triggers. *Linguistic Inquiry*, 25, 407-454, 1994.

[Gigerenzer, 2002] G. Gigerenzer. *Reckoning with risk: Learning to live with uncertainty*. Harmondsworth, UK: Penguin Books, 2002.

[Gigerenzer, 1991] G. Gigerenzer. From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, 98, 254-267, 1991.

[Gigerenzer and Goldstein, 1996] G. Gigerenzer and D. Goldstein. Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–669, 1996.

[Gigerenzer and Hoffrage, 1995] G. Gigerenzer and U. Hoffrage. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684-704, 1995.

[Gigerenzer and Murray, 1987] G. Gigerenzer and D. J. Murray. *Cognition as intuitive statistics*. Hillsdale, NJ: Erlbaum, 1987.

[Gigerenzer *et al.*, 1989] G. Gigerenzer, Z. Swijinck, T. Porter, L. Daston, J. Beatty, and L. Kruger. *The empire of chance*. Cambridge, England: Cambridge University Press, 1989.

[Gigerenzer *et al.*, 1999] G. Gigerenzer, P. Todd, and The ABC Group, eds. *Simple heuristics that make us smart.* Oxford: Oxford University Press, 1999.

[Ginsberg, 1987] M. Ginsberg. *Readings in nonmonotonic reasoning*. Morgan Kaufmann Publishers, 1987.

[Glymour, 1980] C. Glymour. *Theory and evidence*. Princeton: Princeton University Press, 1980.

[Gold, 1967a] E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.

[Gold, 1967b] E. M. Gold. Language identification in the limit. *Information and Control*, 10, 447-474, 1967.

[Gold and Shadlen, 2000] J. I. Gold and M. N. Shadlen. Representation of a perceptual decision in developing oculomotor commands. *Nature*, 404, 390-394, 2000.

[Goldsmith, 2001] J. Goldsmith. Unsupervised learning of the morphology of a natural language. *Computational Linguistics*, 27, 153-198, 2001.

[Goodman, 1951] N. Goodman. *The structure of appearance*. Cambridge, MA: Harvard University Press, 1951.

[Goodman, 1954] N. Goodman. *Fact, fiction, and forecast*. London: The Athlone Press, 1954.

[Gopnik *et al.*, 2004] A. Gopnik, C. Glymour, D. M. Sobel, L. E. Schulz, T. Kushnir, and D. Danks. A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, 111, 1-31, 2004.

[Green and Over, 1997] D. W. Green and D. E. Over. Causal inference, contingency tables and the selection task. *Current Psychology of Cognition*, 16, 459-487, 1997.

[Green and Over, 2000] D. W. Green and D. E. Over. Decision theoretical effects in testing a causal conditional. *Current Psychology of Cognition*, 19, 51-68, 2000.

[Gregory, 1970] R. L. Gregory. *The Intelligent Eye*. London: Weidenfeld and Nicolson, 1970.

[Griffiths and Tennenbaum, 2006] T. L. Griffiths and J. B. Tenenbaum. Optimal predictions in everyday cognition. *Psychological Science*,17, 767-773, 2006.

[Griffiths and Tennenbaum, 2005] T. L. Griffiths and J. B. Tenenbaum. Structure and strength in causal induction. *Cognitive Psychology*, 51, 354-384, 2005.

[Griffiths *et al.*, 2007] T. L. Griffiths, M. Steyvers, and J. B. Tenenbaum. Topics in semantic representation. *Psychological Review*, 114, 211-244, 2007.

[Griffiths *et al.*, 2005] T. L. Griffiths, M. Steyvers, D. M. Blei, and J. B. Tenenbaum. Integrating topics and syntax. *Advances in Neural Information Processing* Systems 17, 2005.

[Griffiths and Steyvers, 2004] T. L. Griffiths and M. Steyvers. Finding scientific topics. *Proceedings of the National Academy of Sciences, 101*, 5228-5235, 2004.

[Grodner and Gibson, 2005] D. Grodner and E. Gibson. Consequences of the serial nature of linguistic input. *Cognitive Science*, 29, 261-291, 2005.

[Hacking, 1975] I. Hacking. *The emergence of probability*. Cambridge, England: Cambridge University Press, 1975.

[Hacking, 1990] I. Hacking. *The taming of chance*. Cambridge, England: Cambridge University Press, 1990.

[Hahn and Nakisa, 2000] U. Hahn and R. Nakisa. German inflection: Single route or dual route? *Cognitive Psychology*, 41, 313-360, 2000.

[Hahn and Oaksford, 2006] U. Hahn and M. Oaksford. A Bayesian approach to informal argument fallacies. *Synthese*, 152, 207-236, 206.

[Hahn and Oaksford, 2007] U. Hahn and M. Oaksford. The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, 114, 704-732, 2007.

[Hahn *et al.*, 2005a] U. Hahn, M. Oaksford, and H. Bayindir. How convinced should we be by negative evidence? In B. Bara, L. Barsalou, and M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, (pp. 887-892), Mahwah, N.J.: Lawrence Erlbaum Associates, 2005.

[Hahn *et al.*, 2005b] U. Hahn, M. Oaksford, and A. Corner. Circular arguments, begging the question and the formalization of argument strength. In A. Russell, T. Honkela, K. Lagus, and M. Pöllä, (Eds.), *Proceedings of AMKLC'05, International Symposium on Adaptive Models of Knowledge, Language and Cognition*, (pp. 34-40), Espoo, Finland, June 2005.

[Hale, 2003] J. Hale. The Information Conveyed by Words in Sentences. *Journal of Psycholinguistic Research*. 32, 101-123, 2003.

[Hamblin, 1970] C. L. Hamblin. *Fallacies*. London: Methuen, 1970.

[Hammond, 1996] K. R. Hammond. *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford: Oxford University Press, 1996.

[Hattori, 2002] M. Hattori. A quantitative model of optimal data selection in Wason's selection task. *Quarterly Journal of Experimental Psychology,*55*A*, 1241-1272. 2002.

[Hay and Baayen, 2005] J. Hay and H. Baayen. Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences*, 9, 342-348, 2005.

[Heit, 2000] E. Heit. Properties of inductive reasoning. *Psychonomic Bulletin and Review, 7*, 569-592, 2000.

[Heit, 1998] E. Heit. A Bayesian analysis of some forms of inductive reasoning. In M. Oaksford and N. Chater (Eds.), *Rational models of cognition* (pp. 248-274). Oxford: Oxford University Press, 1998.

[Heit and Rubinstein, 1994] E. Heit and J. Rubinstein. Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 411-422, 1994.

[Helmholtz, 1910/1962] H. von Helmholtz. *Treatise on Physiological Optics*, Vol. 3, J. P. Southall (Ed. and translation). New York, NY: Dover Publications, 1910/1932.

[Helpel, 1945] C. G. Hempel. Studies in the logic of confirmation. *Mind*, 54, 1-26, 97-121 1945.

[Henle, 1978] M. Henle. Foreword to R. Revlin and R. E. Mayer (Eds.), *Human reasoning*. Washington: Winston, 1978.

[Hertwig *et al.*, 2004] R. Hertwig, G. Barron, E. U. Weber, and I. Erev. Decisions from experience and the effect of rare events in risky choices. *Psychological Science*, 15, 534-539, 2004.

[Hochberg and McAlister, 1953] J. E. Hochberg and E. and McAlister. A quantitative approach to figural "goodness." *Journal of Experimental Psychology, 46*, 361–364, 1953.

[Hogarth and Karelaia, 2005a] R. M. Hogarth and N. Karelaia. Ignoring information in binary choice with continuous variables: When is less "more"? *Journal of Mathematical Psychology, 49*, 115-124, 2005.

[Hogart and Karelaia, 2005b] R. M. Hogarth and N. Karelaia. Simple models for multi-attribute choice with many alternatives: When it does and does not pay to face trade-offs with binary attributes. *Management Science*, 51, 1860-1872, 2005.

[Horning, 1971] J. Horning. A Procedure for Grammatical Inference. *Proceedings of the IFIP Congress 71* (pp. 519-523), Amsterdam: North Holland, 1971.

[Horwich, 1982] P. Horwich. *Probability and Evidence*. New York: Cambridge University Press. 1982.

[Howe and Purves, 2004] C. Q. Howe and D. and Purves. Size contrast and assimilation explained by the statistics of natural scene geometry. *Journal of Cognitive Neuroscience*, 16, 90-102, 2004.

[Howe and Purves, 2005] C. Q. Howe and D. Purves. *Perceiving Geometry: Geometrical Illusions Explained by Natural Scene Statistics*. Berlin: Springer, 2005.

[Howson and Urbach, 1993] C. Howson and P. Urbach. *Scientific reasoning: The Bayesian approach* (2^{nd} edition). La Salle, IL: Open Court, 1993.

[Hutter, 2004] M. Hutter. *Universal Artificial Intelligence: Sequential Decisions Based on Algorithmic Probability*. Berlin: Springer, 2004.

[Inhelder and Piaget, 1955] B. Inhelder and J. Piaget. *De la logique de l'enfant à la logique de l'adolescent*. Paris: Presses Universitaires de France, 1955. (English version: The growth of logical thinking from childhood to adolescence. London: Routledge, 1958).

[Jeffrey, 1965] R. Jeffrey. *The logic of decision*. New York: McGraw Hill, 1965.

[Jeffrey, 1983] R. Jeffrey. *The Logic of decision*. 2nd ed, Chicago, University of Chicago Press, 1983.

[Johnson and Riezler, 2002] M. Johnson and S. Riezler. Statistical models of language learning and use. *Cognitive Science*,26, 239-253, 2002.

[Johnson-Laird, 1983] P. N. Johnson-Laird. *Mental models*. Cambridge, England: Cambridge University Press, 1983.

[Johnson-Laird and Byrne, 2002] P. N. Johnson-Laird and R. M. J. Byrne. Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109,* 646-678, 2002.

[Johnson-Laird and Steedman, 1978] P. N. Johnson-Laird and M. Steedman. The psychology of syllogisms. *Cognitive Psychology, 10,* 64–99, 1978.

[Johnson-Laird *et al.*, 1999] P. N. Johnson-Laird, P. Legrenzi, V. Girotto, M. S. Legrenzi, and J. P. Caverni. Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, 106, 62–88, 1999.

[Johnson-Laird and Byrne, 1991] P. N. Johnson-Laird and R. M. J. Byrne. *Deduction.* Hillsdale, NJ: Lawrence Erlbaum Associates 1991.

[Jurafsky, 1996] D. Jurafsky. A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137–194, 1996.

[Jurafsky, 2003a] D. Jurafsky. Pragmatics and Computational Linguistics. In Laurence R. Horn and Gregory Ward (eds.) Handbook of Pragmatics. Oxford: Blackwell, 2003.

[Jurafsky, 2003b] D. Jurafsky. Probabilistic Modeling in Psycholinguistics: Linguistic Comprehension and Production. In R. Bod, J. Hay, and S. Jannedy, (Eds.), *Probabilistic Linguistics* (pp. 291-320). Cambridge, MA: MIT Press, 2003.

[Kahneman, 2000] D. Kahneman. Preface. In D. Kahneman and A. Tversky, (Eds.), *Choices, values and frames* (pp. ix-xvii). New York: Cambridge University Press and the Russell Sage Foundation, 2000.

[Kahneman and Tversky, 2000] D. Kahneman and A. Tversky, eds. *Choices, values and frames.* New York: Cambridge University Press and the Russell Sage Foundation, 2000.

[Kahneman and Tversky, 1979] D. Kahneman and A. Tversky. Prospect theory: An analysis of decisions under risk. *Econometrica*, 47, 313-327, 1979.

[Kahneman *et al.*, 1982] D. Kahneman, P. Slovic, and A. Tversky, eds. *Judgment under uncertainty: Heuristics and biases.* New York, NY: Cambridge University Press, 1982.

[Kakade and Dayan, 2002] S. Kakade and P. Dayan. Acquisition and extinction in autoshaping. *Psychological Review*, 109, 533-544, 2002.

[Kant, 1787/1961] E. Kant. *Critique of the pure reason.* (trans. N. K. Smith), First Edition, Second Impression. London, England: Macmillan, 1787/1961.

[Kemp and Tenenbaum, 2009] C. Kemp and J. B. Tenenbaum. Structured statistical models of inductive reasoning. *Psychological Review*, 116, 20-58, 2009.

[Kirby, 1994] K. N. Kirby. Probabilities and utilities of fictional outcomes in Wason's four card selection task. *Cognition, 51,* 1-28, 1994.

[Klauer, 1999] K. C. Klauer. On the normative justification for information gain in Wason's selection task. *Psychological Review, 106,* 215-222, 1999.

[Klauer *et al.*, 2006] K. C. Klauer, C. Stahl, and E. Erdfelder. The abstract selection task: An almost comprehensive model. Unpublished manuscript. Albert-Ludwigs-Universität Freiburg, 2006.

[Klavans and Resnik, 1996] J. Klavans and P. Resnik, eds. *The Balancing Act: Combining Symbolic and Statistical Approaches to Language.* MIT Press, Cambridge, MA, 1996.

[Klein and Manning, 2002] D. Klein and C. Manning. A generative constituent-context model for improved grammar induction. In *Proceedings of the 40^{th} Annual Meeting of the ACL*, 2002.

[Klein and Manning, 2002] D. Klein and C. Manning. A generative constituent-context model for improved grammar induction. In *ACL 40*, pages 128–135, 2002.

[Klein and Manning, 2004] D. Klein and C. Manning. Corpus-based induction of syntactic structure: models of sependency and constituency. In *Proceedings of the 42^{nd} Annual Meeting of the ACL*, 2004.

[Knill and Richards, 1996] D. C. Knill and W. A. Richards, eds. *Perception as Bayesian inference*. Cambridge, England: Cambridge University Press, 1996.

[Knill and Saunders, 2003] D. C. Knill and J. A. Saunders. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research*, 43, 2539–2558, 2003.

[Körding and Wolpert, 2006] K. P. Körding and D. M. Wolpert. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10, 319-326, 2006.

[Körding and Wolpert, 2004] K. P. Körding and D. M. Wolpert. Bayesian integration in sensorimotor learning, *Nature*, 427, 244-247, 2004.

[Körding *et al.*, 2004] K. P. Körding, S. P. Ku, and D. Wolpert. Bayesian integration in force estimation. *Journal of Neurophysiology*, 92, 3161-3165, 2004.

[Krebs and Davies, 1996] J. R. Krebs and N. Davies, eds. *Behavioural ecology: An evolutionary approach* (4^{th} edition). Oxford: Blackwell, 1996.

[Kruschke, 2006] J. K. Kruschke. Local Bayesian learning with applications to retrospective revaluation and highlighting. *Psychological Review*, 113, 677-699, 2006.

[Kuhn, 1962] T. Kuhn. *The structure of scientific revolutions*. Chicago: University of Chicago Press, 1962.

[Lakatos, 1970] I. Lakatos. Falsification and the methodology of scientific research programmes. In I. Lakatos, and A. Musgrave (Eds.) *Criticism and the growth of knowledge* (pp. 91-196). Cambridge, England: Cambridge University Press, 1970.

[Laming, 1997] D. Laming. *The measurement of sensation*. Oxford: Oxford University Press. 1997.

[Landaur and Dumais, 1997] T. K. Landauer and S. T. Dumais. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211-240, 1997.

[Lari and Young, 1990] K. Lari and S. Y. Young. The estimation of stochastic context-free grammars using the inside-outside algorithm. *Computer Speech and Language*, 4:35–56, 1990.

[Laudan and Leplin, 1991] L. Laudan and J. Leplin. Empirical equivalence and underdetermination. *Journal of Philosophy*, 88, 449-472, 1991.

[Leeuwenberg, 1969] E. Leeuwenberg. Quantitative specification of information in sequential patterns. *Psychological Review, 76*, 216-220, 1969.

[Leeuwenberg, 1971] E. Leeuwenberg. A perceptual coding language for perceptual and auditory patterns. *American Journal of Psychology*, 84, 307-349. 1971.

[Leeuwenberg and Boselie, 1988] E. Leeuwenberg and E. Boselie. Against the likelihood principle in visual form perception. *Psychological Review,* 95, 485-491, 1988.

[Legate and Yang, 2002] J. A. Legate and C. D. Yang. Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review* 19 (2002), 151–162, 2002.

[Li and Vitanyi, 1997] M. Li and P. M. B. Vitanyi. *An Introduction to Kolmogorov Complexity and its Applications* (*Second Edition*) Springer-Verlag, New York, 1997.

[Lindley, 1956] D. V. Lindley. On a measurement of the information provided by an experiment. *Annals of Mathematical Statistics, 27*, 986-1005, 1956.

[Long and Purves, 2003] F. Long and D. Purves. Natural scene statistics as the universal basis for color context effects. *Proceedings of the National Academy of Science*, 100, 15190-15193, 2003.

[Loomes and Sugden, 1982] G. Loomes amd R. Sugden. Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal, 92,* 805-824, 1982.

[MacDonald *et al.*, 1994] M. C. MacDonald, N. Pearlmutter, and M. S. Seidenberg. The lexical nature of syntactic ambiguity resolution. *Psychological Review,* 101, 676-703, 1994.

[Mach, 1959] E. Mach. *The analysis of sensations and the relation of the physical to the psychical.* New York: Dover Publications, 1959. (Original work published 1914.)

[Mackay, 1992] D. J. C. Mackay. Bayesian interpolation. *Neural Computation*, 4, 415-447. 1992.

[Mackay, 2003] D. J. C. Mackay. *Information theory, inference, and learning algorithms*. Cambridge University Press: Cambridge, 2003.

[Manktelow and Over, 1987] K. I. Manktelow and D. E. . Reasoning and rationality. *Mind and Language*, 2, 199-219, 1987.

[Manktelow *et al.*, 1995] K. I. Manktelow, E. J. Sutherland, and D. E. Over. Probabilistic factors in deontic reasoning. *Thinking and Reasoning*, 1, 201-220, 1995.

[Manktelow and Over, 1991] K. I. Manktelow and D. E. Over. Social roles and utilities in reasoning with deontic conditionals. *Cognition*, 39, 85-105, 1991.

[Manning, 2003] C. Manning. Probabilistic Syntax. In Rens Bod, Jennifer Hay, and Stefanie Jannedy (eds), *Probabilistic Linguistics*, pp. 289-341. Cambridge, MA: MIT Press, 2003.

[Marcus *et al.*, 1999] G. F. Marcus, S. Vijayan, S. Bandi Rao, and P. M. Vishton. Rule learning by seven-month-old infants, *Science*, 283, 77-80, 1999.

[Marcus and Rips, 1979] S. L. Marcus and L. J. Rips. Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior, 18,* 199-223, 1979.

[Marr, 1982] D. Marr. *Vision*. San Francisco, CA: Freeman, 1982.

[Massaro, 1987] D. W. Massaro. *Speech perception by ear and eye*. Hillsdale, NJ: Erlbaum, 1987.

[McCarthy and Hayes, 1969] J. McCarthy and P. J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Michie (Eds.), *Machine intelligence 4*. Edinburgh, Scotland: Edinburgh University Press, 1969.

[McClelland, 1998] J. L. McClelland. Connectionist models and Bayesian inference. In M. Oaksford and N. Chater, (Eds.), *Rational models of cognition* (pp. 21-53). Oxford, England: Oxford University Press, 1998.

[McClelland and Elman, 1986] J. L. McClelland and J. L. Elman. The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86, 1986.

[McDonald and Shillcock, 2003] S. A. McDonald and R. C. Shillcock. Eye movements reveal the on-line computation of lexical probabilities. *Psychological Science, 14,* 648-652, 2003.

[McKenzie, 2004] C. R. M. McKenzie. Framing effects in inference tasks – and why they are normatively defensible. *Memory and Cognition*, 32**,** 874-885, 2004.

[McKenzie and Mikkelsen, 2000] C. R. M. McKenzie and L. A. Mikkelsen. The psychological side of Hempel's paradox of confirmation. *Psychonomic Bulletin and Review*, 7, 360-366, 2000.

[McKenzie and Mikkelsen, 2007] C. R. M. McKenzie and L. A. Mikkelsen. A Bayesian view of covariation assessment. *Cognitive Psychology, 54,* 33-61, 2007.

[McKenzie *et al.*, 2001] C. R. M. McKenzie, V. S. Ferreira, L. A. Mikkelsen, K. J. McDermott, and R. P. Skrable. Do conditional statements target rare events? *Organizational Behavior and Human Decision Processes*, 85, 291-309, 2001.

[McRae *et al.*, 1998] K. McRae, M. J. Spivey-Knowlton, and M. K. Tanenhaus. Modeling the influence of thematic fit (and other constraints) in online sentence comprehension. *Journal of Memory and Language, 38*, 283-312, 1998.

[Medin *et al.*, 1997] D. L. Medin, E. B. Lynch, J. D. Coley, and S. Atran. Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology, 32*, 49-96, 1997.

[Miller, 1956] G. A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for information processing. *Psychological Review*, 63,81–97, 1956.

[Miyazaki *et al.*, 2005] M. Miyazaki, D. Nozaki, and Y. Nakajima. Testing Bayesian models of human coincidence timing. *Journal of Neurophysiology*, 94, 395-399, 2005.

[Monaghan *et al.*, 2007] P. Monaghan, M. Christiansen, and N. Chater. The Phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55, 259-305, 2007.

[Narayanan and Jurafsky, 2002] S. Narayanan and D. Jurafsky. A Bayesian model predicts human parse preference and reading time in sentence processing. In T. G. Dietterich, S. Becker and Z. Ghahramani (Eds.*), Advances in neural information processing systems* (volume 14, pp. 59-65). Cambridge, MA: MIT Press, 2002.

[Nelson, 2005] J. Nelson. Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, 112, 979-999, 2005.

[Newell and Simon, 1972] A. Newell and H. A. Simon. *Human problem solving*. Englewood Cliffs, N.J: Prentice-Hall, 1972.

[Newell *et al.*, 1958] A. Newell, J. C. Shaw, and H. A. Simon. Chess-playing programs and the problem of complexity. *IBM Journal of Research and Development*, 2, 320-25 1958.

[Newstead *et al.*, 1999] S. E. Newstead, S. J. Handley, and E. Buck. Falsifying mental models: Testing the predictions of theories of syllogistic reasoning. *Memory and Cognition*, 27, 344-354, 1999.

[Nickerson, 1996] R. S. Nickerson. Hempel 's paradox and Wason 's selection task: Logical and psychological puzzles of confirmation. *Thinking and Reasoning*, 2, 1-32, 1996.

[Nisbett *et al.*, 1983] R. E. Nisbett, D. H. Krantz, C. Jepson, and Z. Kunda. The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90*, 339-363, 1983.

[Niyogi, 2006] P. Niyogi. *The Computational Nature of Language Learning and Evolution.* Cambridge, MA: MIT Press, 2006.

[Norris, 2006] D. Norris. The Bayesian Reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review*, 113, 327-357, 2006.

[Novick and Cheng, 2004] L. R. Novick and P. W. Cheng. Assessing interactive causal influence. *Psychological Review*, 111, 455-485, 2004.

[Nundy and Purves, 2002] S. Nundy and D. Purves. A probabilistic explanation of brightness scaling. *Proceedings of the National Academy of Sciences*, 99, 14482-14487, 2002.

[Oaksford, 2004a] M. Oaksford. *Conditional inference and constraint satisfaction*: *Reconciling probabilistic and mental models approaches?* Paper presented at the 5^{th} International Conference on Thinking, University of Leuven, Leuven, Belgium, 2004.

[Oaksford, 2004b] M. Oaksford. Reasoning. In N. Braisby and A. Gellatly (Eds.), *Cognitive psychology* (pp. 418-455). Oxford, England: Oxford University Press, 2004.

[Oaksford and Chater, 1991] M. Oaksford and N. Chater. Against logicist cognitive science. *Mind and Language, 6*, 1-38, 1991.

[Oaksford and Chater, 1994] M. Oaksford and N. Chater. A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608-631. 1994.

[Oaksford and Chater, 1996] M. Oaksford and N. Chater. Rational explanation of the selection task. *Psychological Review*, 103, 381-391, 1996.

[Oaksford and Chater, 1998a] M. Oaksford and N. Chater, eds. *Rational models of cognition*, Oxford University Press, Oxford, 1998.

[Oaksford and Chater, 1998b] M. Oaksford and N. Chater. *Rationality in an uncertain world*. Hove, England: Psychology Press, 1998.

[Oaksford and Chater, 2003a] M. Oaksford and N. Chater. Conditional probability and the cognitive science of conditional reasoning. *Mind and Language*, 18, 359-379. 2003.

[Oaksford and Chater, 2003b] M. Oaksford and N. Chater. Optimal data selection: Revision, review and re-evaluation. *Psychonomic Bulletin and Review*, 10, 289-318. 2003.

[Oaksford and Chater, 2007] M. Oaksford and N. Chater. *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press 2007.

[Oaksford and Chater, 2008] M. Oaksford and N. Chater. Probability logic and the *Modus Ponens-Modus Tollens* asymmetry in conditional inference. In N. Chater, and M. Oaksford (Eds.), *The probabilistic mind*: *Prospects for Bayesian cognitive science* (pp. 97-120). Oxford: Oxford University Press, 2008.

[Oaksford and Hahn, 2004] M. Oaksford and U. Hahn. A Bayesian analysis of the argument from ignorance. *Canadian Journal of Experimental Psychology*, 58, 75-85, 2004.

[Oaksford and Hahn, 2007] M. Oaksford and U. Hahn. Induction, deduction and argument strength in human reasoning and argumentation. In A. Feeney, and E. Heit (Eds.), *Inductive reasoning* (pp. 269-301). Cambridge: Cambridge University Press, 2007.

[Oaksford and Moussakowski, 2004] M. Oaksford and M. Moussakowski. Negations and natural sampling in data selection: Ecological vs. heuristic explanations of matching bias. *Memory and Cognition*, 32, 570-581, 2004.

[Oaksford and Stenning, 1992] M. Oaksford and K. Stenning. Reasoning with conditionals containing negated constituents. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 835-854, 1992.

[Oaksford and Wakefield, 2003] M. Oaksford and M. Wakefield. Data selection and natural sampling: Probabilities do matter. *Memory and Cognition*, 31, 143-154, 2003.

[Oaksford *et al.*, 1999] M. Oaksford, N. Chater, and B. Grainger. Probabilistic effects in data selection. *Thinking and Reasoning, 5,* 193-244, 1999.

[Oaksford *et al.*, 2000] M. Oaksford, N. Chater, and J. Larkin. Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26,* 883-889, 2000.

[Oaksford *et al.*, 1997] M. Oaksford, N. Chater, B. Grainger, and J. Larkin. Optimal data selection in the reduced array selection task (RAST). *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 441-458, 1997.

[Oaksford *et al.*, 2002] M. Oaksford, L. Roberts, and N. Chater. Relative informativeness of quantifiers used in syllogistic reasoning. *Memory and Cognition*, 30, 138-149, 2002.

[Oberauer and Wilhelm, 2003] K. Oberauer and O. Wilhelm. The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 29, 680-693, 2003.

[Oberauer, 2006] K. Oberauer. Reasoning with conditionals: A test of formal models of four theories. *Cognitive Psychology*, 53, 238-283. 2006.

[Oberauer *et al.*, 2004] K. Oberauer, A., Weidenfeld, and R. Hörnig. Logical reasoning and probabilities: A comprehensive test of Oaksford and Chater (2001). *Psychonomic Bulletin and Review*, 11, 521-527, 2004.

[Oberauer *et al.*, 1999] K. Oberauer, O. Wilhelm, and R. R. Dias. Bayesian rationality for the Wason selection task? A test of optimal data selection theory. *Thinking and Reasoning*, 5, 115-144, 1999.

[Olivers *et al.*, 2004] C. L. N. Olivers, N. Chater, and D. G. Watson. Holography does not account for goodness: A critique of van der Helm and Leeuwenberg (1996). *Psychological Review, 111,* 242–260, 2004.

[Osherson *et al.*, 1990] D. N. Osherson, E. E. Smith, O. Wilkie, A. Lopez, and E. Shafir. Category-based induction. *Psychological Review, 97*, 185-200, 1990.

[Over and Evans, 1994] D. E. Over and J. St. B. T. Evans. Hits and misses: Kirby on the selection
task. *Cognition*, 52, 235–243, 1994.

[Over and Jessop, 1998] D. E. Over and A. Jessop. Rational analysis of causal conditionals and the selection task. In M. Oaksford and N.Chater (Eds.), *Rational Models of Cognition* (pp. 399–414). Oxford, England: Oxford University Press. 1998.

[Over *et al.*, 2007] D. E. Over, C. Hadjichristidis, J. St. B. T. Evans, S. J. Handley, and S. A. Sloman. The psychology of causal conditionals. *Cognitive Psychology*, 54, 62-97, 1007.

[Pearce, 1997] J. M. Pearce. *Animal Learning and Cognition: An Introduction.* Hove: Psychology Press, 1997.

[Pearl, 1988] J. Pearl. *Probabilistic reasoning in intelligent systems.* San Mateo: Morgan Kaufmann, 1998.

[Pearl, 2000] J. Pearl. *Causality: Models, reasoning and inference.* Cambridge, England: Cambridge University Press, 2000.

[Perham and Oaksford, 2005] N. Perham and M. Oaksford. Deontic reasoning with emotional content: Evolutionary psychology or decision theory? *Cognitive Science*, 29, 681-718, 2005.

[Pfeifer and Kleiter, 2005] N. Pfeifer and G. D. Kleiter. Toward a mental probability logic. *Psychologica Belgica*, 45, 71-99, 2005.

[Pickering *et al.*, 2000] M. J. Pickering, M. J. Traxler, and M. W. Crocker. Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language* 43, 447-475. 2000.

[Pierrehumbert, 2001] J. Pierrehumbert. Stochastic phonology. *GLOT*, 5(6), 1-13, 2001.

[Pinker, 1979] S. Pinker. Formal models of language learning. *Cognition*, 7, 217-283, 1979.

[Pinker, 1999] S. Pinker. *Words and rules: The ingredients of language.* New York: Basic Books. 1999.

[Politzer and Braine, 1991] G. Politzer and M. D. Braine. Responses to inconsistent premises cannot count as suppression of valid inferences. *Cognition,* 38, 103-108, 1991.

[Pomerantz and Kubovy, 1986] J. R. Pomerantz and M. Kubovy. Theoretical approaches to perceptual organization: simplicity and likelihood principles. In: K.R. Boff, L. Kaufnam and J. P.Thomas (Eds.) *Handbook of perception and human performance, Volume II: Cognitive processes and performance*. (pp.36:1-45). New York: Wiley, 1986.

[Popper, 1935/1959] K. Popper. *The logic of scientific discovery.* Basic Books, New York, 1935/1959.

[Pothos and Chater, 2002] E. Pothos and N. Chater. A simplicity principle in unsupervised human categorization. *Cognitive Science*, 26, 303-343, 2002.

[Pullum and Scholz, 2002] G. Pullum and B. Scholz. Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 19, 9–50, 2002.

[Putnam, 1974] H. Putnam. The 'corroboration' of theories", in A. Schilpp (Ed.), *The Philosophy of Karl Popper* (Vol. 2), La Salle, IL: Open Court, 1974.

[Pylyshyn, 1987] Z. Pylyshyn, ed. *The robot's dilemma: The frame problem in artificial intelligence.* Norwood, NJ: Ablex, 1987.

[Quiggin, 1993] J. Quiggin. *Generalized expected utility theory: The rank-dependent model.* Norwell, MA: Kluwer Academic Publishers, 1993.

[Quine, 1953] W. V. O. Quine. *From a logical point of view*, Cambridge, MA: Harvard University Press. 1953.

[Rabin, 2000] M. Rabin. Diminishing Marginal Utility of Wealth Cannot Explain Risk Aversion. In D. Kahneman and A. Tversky (Eds.) *Choices, Values, and Frames* (pp. 202-208). New York: Cambridge University Press, 2000.

[Ramachandran, 1990] V. S. Ramachandran. The Utilitarian Theory of Perception. In C. Blakemore (Ed.) *Vision: Coding and Efficiency* (pp. 346–360). Cambridge: Cambridge University Press, 1990.

[Ramsey, 1931/1990] F. P. Ramsey. *The foundations of mathematics and other logical essays.* London: Routledge and Kegan Paul, 1931/1990.

[Redington and Chater, 1998] M. Redington and N. Chater. Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes, 13,* 129-191, 1998.

[Redington *et al.*, 1998a] M. Redington, N. Chater, and S. Finch. Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 425-469, 1998.

[Redington *et al.*, 1998b] M. Redington, N. Chater, and S. Finch. Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 425-469, 1998.

[Reiter, 1980] R. Reiter. A logic for default reasoning, *Artificial Intelligence*,13, 81–132, 1980.

[Restle, 1970] E. Restle. Theory of serial pattern learning: Structural trees. *Psychological Review*,77, 481-495, 1970.

[Rieke *et al.*, 1997] F. Rieke, R. De Ruyter Van Steveninck, D. Warland, and W. Bialek. *Spikes: Exploring the neural code*. Cambridge, MA: MIT Press, 1997.

[Rips, 1975] L. J. Rips. Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14*, 665-681, 1975.

[Rips, 1983] L. J. Rips. Cognitive processes in propositional reasoning. *Psychological Review, 90*, 38-71, 1983.

[Rips, 1994] L. J. Rips. *The psychology of proof.* Cambridge, MA: MIT Press 1994.

[Rips, 2001] L. J. Rips. Two kinds of reasoning. *Psychological Science*, 12, 129-134, 2001.

[Rissanen, 1987] J. Rissanen. Stochastic complexity. *Journal of the Royal Statistical Society, Series B,* 49, 223-239, 1987.

[Rissanen, 1996] J. Rissanen. Fisher information and stochastic complexity. *IEEE Transactions of Information Theory*, 42, 40-47, 1996.

[Roberts and Pashler, 2000] S. Roberts and H. Pashler. How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107,* 358-367, 2000.

[Rock, 1983] I. Rock. *The logic of perception*. Cambridge, MA: MIT Press, 1983.

[Rosch, 1975] E. Rosch. Cognitive representation of semantic categories. *Journal of experimental psychology: General*, 104, 192-233, 1975.

[Rubenstein, 1998] A. Rubenstein. *Modeling bounded rationality.* Cambridge, MA: MIT Press, 1998.

[Rumelhart *et al.*, 1986] D. E. Rumelhart, P. Smolensky, J. L. McClelland, and G. E. Hinton. Schemata and sequential thought processes in PDP models, in: J.McClelland and D.Rumelhart (Eds) *Parallel distributed processing: Explorations in the microstructure of cognition Vol 2: Psychological and biological models* (MIT Press), 1986.

[Saumuelson and Zeckhauser, 1988] W. F. Samuelson and R. J. Zeckhauser. Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1, 7-59, 1988.

[Saunders and Knill, 2004] J. A. Saunders and D. C. Knill. Visual feedback control of hand movements. *Journal of Neuroscience*, 24, 3223-3234, 2004.

[Savage, 1954] L. J. Savage. *The Foundations of Statistics*. New York, NY: Wiley, 1954.

[Schooler and Anderson, 1997] L. J. Schooler and J. R. Anderson. The role of process in the rational analysis of memory. *Cognitive Psychology*, 32, 219-250, 1997.

[Schrater and Kersten, 2000] P. R. Schrater and D. Kersten. How optimal depth cue integration depends on the task. *International Journal of Computer Vision*, 40, 71-89, 2000.

[Schroyens and Schaeken, 2003] W. Schroyens and W. Schaeken. A critique of Oaksford, Chater and Larkin's (2000) conditional probability model of conditional reasoning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 29,* 140-149, 2003.

[Schütze, 1995] H. Schütze. Distributional part-of-speech tagging. In *Proc. of 7th Conference of the European Chapter of the Association for Computational Linguistics,* pp. 141-148, 1995.

[Schwarz, 1978] G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464, 1978.

[Seidenberg and Elman, 1999] M. S. Seidenberg and J. L. Elman. Do infants learn grammar with algebra or statistics? *Science*, 284, 434-435, 1999.

[Seidenberg, 1997] M. S. Seidenberg. Language acquisition and use: learning and applying probabilistic constraints. *Science*, 275, 1599-1603, 1997

[Shanks, 1995] D. R. Shanks. *The psychology of associative learning*. Cambridge: Cambridge University Press, 1995.

[Shannon, 1951] C. E. Shannon. *Prediction and entropy of printed English*. Bell System Technical Journal, 30(1):50-64, January 1951.

[Shannon and Weaver, 1949] C. E. Shannon and W. Weaver. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.

[Shiffrin and Steyvers, 1998] R. M. Shiffrin and M. Steyvers. The effectiveness of retrieval from memory. In M. Oaksford and N. Chater (Eds.), *Rational Models of Cognition* (pp. 73-95) Oxford: Oxford University Press, 1998.

[Simon, 1957] H. A. Simon. *Models of man*, New York, NY: Wiley, 1957.

[Simpson, 1951] E. H. Simpson. The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Ser, B, 13,* 238-241, 1951.

[Skyrms, 1986] B. Skyrms. *Choice and chance: An introduction to inductive logic*. Belmont, California: Wadsworth, 1986.

[Sloman, 1993] S. A. Sloman. Feature-based induction. *Cognitive Psychology, 25*, 231-280, 1993.

[Sloman and Lagnado, 2005] S. A. Sloman and D. Lagnado. Do we "do"? *Cognitive Science*, 29, 5-39, 2005.

[Smolensky and Legendre, 2006] P. Smolensky and G. Legendre. *The harmonic mind* (2 Vols). Cambridge, MA: MIT Press, 2006.

[Snippe *et al.*, 2000] H. P. Snippe, L. Poot, and J. H. van Hateren. A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds. *Visual Neuroscience 17*, 449-462, 2000.

[Sobel, 2004] J. H. Sobel. *Probable modus ponens and modus tollens and updating on uncertain evidence.* Unpublished manuscript, Department of Philosophy, University of Toronto, Scarborough, 2004.

[Sober, 2002] E. Sober. Intelligent design and probability reasoning. *International Journal for Philosophy of Religion, 52,* 65-80, 2002.

[Stanovich and West, 2000] K. E. Stanovich and R. F. West. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645-665, 2000.

[Stephens and Krebs, 1986] D. W. Stephens and J. R. Krebs. *Foraging theory.* Princeton, NJ: Princeton University Press, 1986.

[Stewart and Simpson, 2008] N. Stewart and K. Simpson. A decision-by-sampling account of decision under risk. In N. Chater and M. Oaksford (Eds.) *The probabilistic mind*: *Prospect for Bayesian cognitive science* (pp. . Oxford: Oxford University Press, 2008.

[Stewart *et al.*, 2005] N. Stewart, G. D. A. Brown, and N. Chater. Absolute identification by relative judgment. *Psychological Review*, 112, 881-911, 2005.

[Stewart *et al.*, 2006] N. Stewart, N. Chater, and G. D. A. Brown. Decision by sampling. *Cognitive Psychology*, 53, 1-26, 2006.

[Stewart *et al.*, 2003] N. Stewart, N. Chater, H. P. Stott, and S. Reimers. Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General*, 132, 23-46, 2003.

[Swier and Stevenson, 2005] R. Swier and S. Stevenson. Exploiting a Verb Lexicon in Automatic Semantic Role Labelling. *Proceedings of the Joint Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing* (HLT/EMNLP-05), 2005.

[Tanenhaus *et al.*, 1995] M. K. Tanenhaus, M. J. Spivey-Knowlton, K. M. Eberhard, and J. E. Sedivy. Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 632-634, 1995.

[Taplin, 1971] J. E. Taplin. Reasoning with conditional sentences. *Journal of Verbal Learning and Verbal Behavior, 10,* 219-225, 1971.

[Tenenbaum and Griffiths, 2001] J. B. Tenenbaum and T. L. Griffiths. Generalization, similarity, and Bayesia inference. *Behavioral and Brain Sciences*, 24, 629-641, 2001.

[Tenenbaum *et al.*, 2006] J. B. Tenenbaum, T. L. Griffiths, and C. Kemp. Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences, 10*, 309-318, 2006.

[Thaler, 1985] R. Thaler. Mental accounting and consumer choice. *Marketing Science*, 4, 199-214, 1985.

[Todorov and Jordon, 2002] E. Todorov and M. I. Jordon. Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, 5, 1226-1235, 2002.

[Todorov, 2004] E. Todorov. Optimality principles in sensorimotor control. *Nature Neuroscience*, 7, 907-915, 2004.

[Tomasello, 2003] M. Tomasello. Constructing a Language: A Usage-Based Theory of Language Acquisition. Cambridge, MA: Harvard University Press, 2003.

[Toutanova *et al.*, 2005a] K. Toutanova, C. Manning, D. Flickinger, and S. Oepen. Stochastic HPSG Parse Disambiguation using the Redwoods Corpus. *Research on Language and Computation*, 3, 83-105, 2005.

[Trommershäuser *et al.*, 2006] J. Trommershäuser, M. S. Landy, and L. T. Maloney. Humans rapidly estimate expected gain in movement planning. *Psychological Science, 11*, 981-988, 2006.

[Tu *et al.*, 2005] Z. Tu, X. Chen, A. L. Yuille, and S.-C. Zhu. Image parsing: Unifying segmentation detection and recognition. *International Journal of Computer Vision*, 2, 113-140, 2005.

[Tversky and Kahneman, 1992] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297-323, 1992.

[van der Helm and Leeuwenberg, 1996] P. A. van der Helm and E. L. J. Leeuwenberg. Goodness of visual regularities: A nontransformational approach. *Psychological Review,* 103,429–456, 1996.

[Verschueren *et al.*, 2005] N. Verschueren, W. Schaeken, and G. d'Ydewalle. A dual-process specification of causal conditional reasoning. *Thinking and Reasoning, 11,* 278-293, 2005.

[von Mises, 1957] R. von Mises. *Probability, statistics and truth* (Revised English Edition). New York, NY: Macmillan, 1957.

[Wagner, 2004] C. G. Wagner. Modus tollens probabilized. *British Journal for Philosophy of Science*, 55, 747-753, 2004.

[Wallace and Freeman, 1987] C. S. Wallace and P. R. Freeman. Estimation and inference by compact coding. *Journal of the Royal Statistical Society*, *Series B, 49*, 240-251, 1987.

[Wallach and O'Connell, 1953] H. Wallach and D. N. O'Connell. The kinetic depth effect. *Journal of Experimental Psychology*, 45, 205-217, 1953.

[Wason and Johnson-Laird, 1972] P. C. Wason and P. M. Johnson-Laird. *Psychology of Reasoning: Structure and Content.* London: Batsford 192.

[Wason, 1960] P. C. Wason. On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140, 1960.

[Wason, 1968] P. C. Wason. Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*, 273-281 1968.

[Weiss, 1997] Y. Weiss. Interpreting images by propagating Bayesian beliefs. In M.C. Mozer, M. I. Jordan and. T. Petsche (Ed.), *Advances in Neural Information Processing Systems 9* (pp. 908-915). Cambridge MA: MIT Press, 1997.

[Xu and Tenenbaum, 2007] F. Xu and J. B. Tenenbaum. Word learning as Bayesian inference. *Psychological Review*, 114, 245-272, 2007.

[Yama, 2001] H. Yama. Matching versus optimal data selection in the Wason selection task. *Thinking and Reasoning, 7,* 295-311, 2001.

[Yuille and Kersten, 2006] A. Yuille and D. Kersten. Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences*, 10, 301-308, 2006.

[Zeelenberg *et al.*, 2000] M. Zeelenberg, W. W. Van Dijk, A. S. R. Manstead, and J. van der Pligt. On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition and Emotion, 14*, 521-541, 2000.

[Zettlemoyer and Collins, 2005] L. S. Zettlemoyer and M. Collins. Learning to map sentences to logical form: Structured classification with probabilistic categorical grammars. In *Proceedings of the Twenty First Conference on Uncertainty in Artificial Intelligence* (UAI-05), 2005.

[Zhai and Lafferty, 2001] C. Zhai and J. Lafferty. Document language models, query models, and risk minimization for information retrieval. In W. Croft, D. Harper, D. Kraft and J. Zobel, (Eds.) *SIGIR Conference on Research and Development in Information Retrieval* (pp. 111–119). New York, NY: ACM Press, 2001.