The uncertain reasoner: Bayes, logic, and rationality

Mike Oaksford and Nick Chater

Behavioral and Brain Sciences / Volume 32 / Issue 01 / February 2009, pp 105 - 120
DOI: 10.1017/S0140525X0900051X, Published online: 12 February 2009

Link to this article: http://journals.cambridge.org/abstract_S0140525X0900051X

How to cite this article:

Request Permissions: Click here
Authors’ Response

The uncertain reasoner: Bayes, logic, and rationality

doi:10.1017/S0140525X0900051X

Mike Oaksford\textsuperscript{a} and Nick Chater\textsuperscript{b}

\textsuperscript{a}School of Psychology, Birkbeck College London, London, WC1E 7HX, United Kingdom; \textsuperscript{b}Division of Psychology and Language Sciences & ESRC Centre for Economic Learning and Social Evolution, University College London, London, WC1E 6BT, United Kingdom.

mike.oaksford@bbk.ac.uk
n.chater@ucl.ac.uk
www.bbk.ac.uk/psyc/staff/academic/oaksford
www.psychol.ucl.ac.uk/people/profiles/chater_nick.htm

Abstract: Human cognition requires coping with a complex and uncertain world. This suggests that dealing with uncertainty may be the central challenge for human reasoning. In Bayesian Rationality we argue that probability theory, the calculus of uncertainty, is the right framework in which to understand everyday reasoning. We also argue that probability theory explains behavior, even on experimental tasks that have been designed to probe people’s logical reasoning abilities. Most commentators agree on the centrality of uncertainty; some suggest that there is a residual role for logic in understanding reasoning; and others put forward alternative formalisms for uncertain reasoning, or raise specific technical, methodological, or empirical challenges. In responding to these points, we aim to clarify the scope and limits of probability and logic in cognitive science; explore the meaning of the “rational” explanation of cognition; and re-evaluate the empirical case for Bayesian rationality.

R1. Introduction

Bayesian Rationality (Oaksford & Chater 2007, henceforth BR) proposed that human reasoning should be understood in probabilistic, not logical, terms. In Part I, we discussed arguments from the philosophy of science, artificial intelligence, and cognitive psychology, which indicate that the vast majority of cognitive problems (outside mathematics) involve uncertain, rather than deductively certain, reasoning. Moreover, we argued that probability theory (the calculus for uncertain reasoning) is a more plausible framework than logic (the calculus for certain reasoning) for modeling both cognition in general, and commonsense reasoning in particular. In Part II, we considered a strong test of this approach, asking whether the probabilistic framework can capture human reasoning performance even on paradigmatically “logical” tasks, such as syllogistic reasoning or conditional inference.

The structure of this response is as follows. In section R2, we reflect on the ubiquity of uncertainty and address the theoretical attempts to preserve logic as a separate and core reasoning process. In section R3, we compare and evaluate Bayesian and logic-based approaches to human reasoning about uncertainty. Section R4 focuses on the methodology of rational analysis (Anderson 1990; 1991a; Oaksford & Chater 1995b) and its relationship to more traditional algorithmic and neuroscientific approaches. Section R5 discusses a variety of specific issues in the empirical data from the psychology of reasoning, and the modeling of that data. Finally, section R6 concludes the case for a “Bayesian turn” in the brain and cognitive sciences in general, and for the understanding of human reasoning in particular.

R2. The ubiquity of uncertainty: Distinctions that might preserve logic

Many commentators suggest ways to preserve a role for logic as a separate and core component in an account of human reasoning, despite the challenge provided by uncertainty (Allott & Uchida, Evans, Politzer & Bonnefon). We argue that logic does have an important role in modeling cognition; but we argue against the existence of cognitive processes dedicated to logical reasoning.

R2.1. Rationality 1 versus Rationality 2

Evans suggests that a distinction should be drawn between two types of rationality (Evans & Over 1996a). Rationality 1 relates to implicit, possibly associative, processes, operating over world knowledge, which Evans also terms “ecological rationality.” This type of rationality arises from System 1 in Evans and Over’s (2004) Dual Process Theory (see also Evans & Frankish, in press; Sloman 1996; Stanovich & West 2000). Rationality 2 involves explicitly following normative rules, and is the type of rationality achieved by Evans and Over’s (2004) System 2. System 2 processes are logical, rule-governed, and conscious. Moreover, Evans has argued for a crucial asymmetry between the systems. It requires cognitive effort to ignore System 1, and to use System 2 for logical inference: that is, to infer only what follows from the structure of the given premises.

The fundamental problem with this Dual Process view is that these two systems must interact – and if the systems obey fundamentally different principles, it is not clear how this is possible. Consider the familiar example of inferring that Tweety flies from the general claim that birds fly and the fact that Tweety is a bird. On the Dual Process view, this inference could be drawn logically from the premises given by System 2, from the assumption that birds fly is a true universal generalization; System 1, by contrast, might tentatively draw this conclusion by defeasible, associative processes, drawing on general knowledge. But a lack of synchrony between the two systems, presumed to operate by different rational standards, threatens to cause inferential chaos. Consider, for example, what happens if we consider the possibility that Tweety is an ostrich. If System 2 works according to logical principles, the clash of two rules threatens contradiction: we know that birds fly, but that ostriches do not. To escape contradiction, one of the premises must be rejected: most naturally, birds fly will be rejected as false. But we now have two unpalatable possibilities. On the one hand, suppose that this retraction is not transferred to general knowledge and hence is not assimilated by System 1. Then the two systems will have contradictory beliefs (moreover, if System 2 reasoning cannot modify general knowledge, its purpose seems unclear). On the
other hand, if birds fly is retracted from world knowledge, along with other defeasible generalizations, then almost all of general knowledge will be stripped away (as BR notes, generalizations outside mathematics are typically defeasible), leading System 1 into inferential paralysis.

The centrality of logic for a putative System 2 is also brought into doubt by considering that one of its main functions is to consciously propose and evaluate arguments. Yet, argumentation, that is, the attempt to persuade oneself or others of a controversial proposition, is uniformly agreed not to be a matter of formal logic (Walton 1989), although aspects of argumentation may naturally be modeled using probability theory (Hahn & Oaksford 2007). Thus, perhaps the core human activity for which a logic-based System 2 is invoked may, ironically, be better explained in probabilistic terms.

People can, of course, be trained to ignore some aspects of linguistic input, and concentrate on others – for example, in the extreme, they can learn to translate natural language statements into predicate logic (ignoring further aspects of their content) and employ logical methods to determine what follows. But, for the psychology of reasoning, this observation is no more significant than the fact that people can learn the rules of chess and ignore most of the visual features of the pieces, the board, or indeed, their surroundings. Conscious application of logical principles is a learned skill built on top of non-logical machinery (and, indeed, is highly effortful, even for logicians); it does not involve, we suggest, tapping in to some underlying logical “engine.”

It is this conscious application of logical concepts (and related notions from mathematics, philosophy, and computer science) that underpins, we suggest, the small but significant correlation between “logical” performance on some reasoning tasks (e.g., selecting the p and not-q cards, in Wason’s selection task) and IQ (Stanovich & West 2000). Logical reasoning is a late and cognitively challenging innovation, rather than a core component of our mental machinery.

Evans also expresses disappointment that we do not address individual differences (Stanovich 2008), which have been viewed as supporting a Dual Process account. But from the present perspective, individual differences concerning the application of learned logical rules are no different from individual differences in chess playing – that is, neither are directly relevant to the question of whether there are single or multiple reasoning systems. Indeed, we suggest that individual differences provide no stronger evidence that cognition involves core logical competence, than that cognition involves core chess-playing competence.

It may turn out, indeed, that there is no real incompatibility between Stanovich’s account and ours. In particular, the distinction Stanovich draws between control processes and other autonomous systems is a distinction common to all theories of cognition (see Oaksford & Chater, in press). But as Kowalski’s (1979) classic equation, “Algorithm = Logic + Control,” reminds us, logic and control processes are very different (see, e.g., Anderson 1983). Hence, Stanovich may not really be committed to anything like Evans’ logically competent System 2. (A further complication is that a distinction between processes of logic and control is now reflected in Evans [2007], who moots the possibility of a tri-process theory.)

**R2.2. The split between semantics and pragmatics**

Griese’s (1975) theory of conversational implicature originally attempted to split off a “striped down” logic-based natural language semantics, from the complex, knowledge-rich processes of pragmatic interpretation involved in inferring a speaker’s intentions. In this way, he aimed to retain a logical core to semantics, despite apparently striking and ubiquitous clashes between the dictates of formal logic and people’s intuitions about meaning and inference.

Within this type of framework, Allott & Uchida attempt to preserve the truth of potentially defeasible conditions (if it’s a bird, then it flies, or, as above, birds fly) despite the ready availability of counterexamples. They suggest that this conditional is true in one model, but not in the model that is considered when an additional premise giving a counterexample is added (e.g., when we consider the possibility that Tweety is an ostrich). But in classical logic, only an inference that holds in all models is deductively valid, by definition. Thus, accepting that this inference holds only in some models implies accepting that the inference is uncertain (contra, e.g., O’Brien). Indeed, in BR, we argue uncertainty is ubiquitous in human reasoning; outside mathematics, deductive reasoning, which guarantees the truth of a conclusion given the premises, is, to a first approximation, never observed.

Moreover, understanding reasoning involves working out pragmatic details about what default background assumptions are applicable in reasoning. Thus, for example, our accounts of specific reasoning phenomena, across conditional reasoning, the selection task, and syllogistic reasoning, involve default assumptions about the environment, for example, what is rare and what is common (cf. McKenzie; McKenzie et al. 2001) and when states are likely to be independent or conditionally independent. In this light, we agree with Stemming & van Lambalgen’s claim that “pure” Bayesian analysis, working from the premises alone, cannot capture suppression effects in conditional reasoning (see sect. R3.6) – we view this as illustrating the knowledge-rich character of reasoning, rather than challenging a Bayesian account.

The ubiquity of uncertain, knowledge-rich inference, argues for an alternative to invoking the semantics/pragmatics distinction to maintain a logical semantics for natural language: namely, that natural language semantics may be probabilistic “all the way down.” Experiments in the psychology of reasoning, as reviewed in BR, find little support for the existence of a level of logic-based representation or inference. BR proposes a starting point for a probabilistic semantics: If p then q conditionals are assumed to express that the conditional probability P(q|p) is high (following Adams 1975; 1998; Bennett 2003; and Edgington 1995, among others); the quantifiers Some, Few, Most, All are similarly assumed to express constraints on probabilities (e.g., Some A are B is rendered as P(A, B) > 0; Most A are B claims that P(B|A) is high)). Switching from a logical to a probabilistic semantics provides, we argue, a better fit with patterns of human reasoning. Of course, it remains possible that a logical core
interpretation might be maintained – but it seems theoretically unparsimonious to do so (Edgington 1995).

A shift from a logical to a probabilistic semantics for aspects of natural language may also allow a more integrated account of semantics and pragmatics. Indeed, McKenzie (e.g., Sher & McKenzie 2006) has powerfully demonstrated the importance of pragmatic factors, even within a purely probabilistic framework (but see, Hilton et al. 2005). Nonetheless, the core insight of Grice’s program remains: that splitting apart semantic factors (concerning meaning) and pragmatic factors (concerning inferences about speaker intentions) is a prerequisite for constructing a tractable semantic theory, whether that theory be based on logic (as Allott & Uchida argue) or probability (as BR proposes).

R2.3. Proof and uncertainty and structure and strength

Politzer & Bonnefon argue that a key element missing from a purely probabilistic account is how premises can be used to construct proofs to derive conclusions. Thus, they argue that the probabilistic account allows the *evaluation* of the strength of the relationship between premises and conclusion, but not how the conclusion is *generated* in the first place. Note, though, that both logic and probability are theories of the nature of inferential relationships between propositions (Harman 1986). Neither specify how reasoning should be carried out, let alone how interesting conclusions should be generated. Moreover, for both logic and probability, a range of algorithms have been developed which can both evaluate given conclusions, and generate new conclusions (e.g., logic programming and Bayesian networks). From both perspectives, any set of information potentially generates an infinite set of possible conclusions; so that an immediate question is: What counts as an *interesting* conclusion? A natural suggestion from the probabilistic point of view is that conclusions with a low prior probability are, other things being equal, more surprising and hence more interesting (as employed in the account of syllogistic reasoning described in BR), although interesting logic-based measures of semantic information content have also been proposed (Johnson-Laird 1983).

More generally, the probabilistic approach is just as able as logic-based approaches to serve as the basis for algorithmic models of thought. For example, Oaksford & Chater (in press) use a constraint satisfaction neural network implementation of the probabilistic approach. The links in the network captures the conditional and default assumptions about *structural* relations between variables (in the causal context, involving alternative causes and defeaters); and the *strength* of each link is captured by a weight. A similar distinction between structure and strength has been invoked in causal reasoning using Bayesian networks (Griffiths & Tenenbaum 2005) and applied in Hahn and Oaksford’s (2007) probabilistic account of argumentation.

R3. Logic, probability, and the challenge of uncertain reasoning?

In this section, we consider whether, as some commentators suggest, we have mischaracterized the scope of logic or chosen the wrong alternative calculus in order to reason about uncertainty. We deal with logic and probability in turn.

R3.1. How are logic and probability related?

Pfeifer & Kleiter observe that probability theory already includes classical propositional logic as a special case. Thus, one way of understanding the approach outlined in BR is as enriching conventional logic to give an *inductive* logic – a system of logic that extends deduction to less-than-certain inferences (Hawthorn 2008). To a good approximation, modern inductive logic just is Bayesian probability (Chater et al., in press; Earman 1992), with some additional discussion of the measure of the confirmation relation (see later discussion of Poletiek and Nelson). Since Carnap (1950), this Bayesian inductive logic includes classical logic – if a statement has a probability of 1, then any logical consequence of that statement also has a probability of 1. Similarly, if a statement has an implication with a probability of 0, then that statement has a probability of 0 (note, however, that probability theory does not readily represent the *internal* structure of atomic propositions, and has no general theory of, for example, quantification or modality). The Bayesian inductive perspective is required not because classic logic is incorrect, but because, outside mathematics, it rarely, if ever, applies (Oaksford & Hahn 2007) – inferential relations between propositions are relentlessly uncertain (Jeffrey 1967).

R3.2. Is relevance relevant?

O’Brien proposes a different enrichment of logic, drawing on his important work with Braine on mental logics (Braine & O’Brien 1991), which aims to capture a notion of *relevance* between antecedent and consequent (i.e., so that conditionals such as *if 2 is odd, then the sky is purple* are no longer automatically true, just in virtue of the false antecedent). Thus, Braine and O’Brien’s work aims to go beyond the material conditional, which BR ascribed to mental logic as a whole (e.g., Rips 1994).

Adding a condition of relevance, while potentially important, does not help deal with the problem of uncertain reasoning, however. Indeed, O’Brien’s account of conditionals is, instead, a strictly deductive version of the Ramsey test (like, e.g., Gärdenfors 1986) – conditionals are only asserted if the consequent, q, follows with certainty from the antecedent p (and background knowledge). Thus, Braine and O’Brien’s (1991) logical interpretation of the conditional suffers the same fundamental problem as material implication: an inability to capture the fact that generalizations outside mathematics are inevitably uncertain.

Moreover, despite Braine and O’Brien’s intentions, their system does not seem to enforce relevance between antecedent and consequent, either. The introduction rule for *if p then q*, used by O’Brien, and described in Braine and O’Brien (1991), states that *if p then q* can be inferred if q follows from the supposition of p together with background knowledge, B. If we know p is false (i.e., background knowledge B implies *not-p*), then supposing p and B implies p & *not-p*, which is a contradiction, from which any conclusion follows – including q. So conditionals
such as if 2 is odd, then the sky is purple can be asserted, after all. Similarly, any conditional whose conclusion is known to be true (i.e., B implies q) will automatically meet the condition that p & B implies q (because this is a monotonic logic – adding premises can never remove conclusions). Hence, conditionals such as if the sky is purple, then 2 is even, will also be asserted – again violating intuitions of relevance.

R3.3. Uncertain reasoning via nonmonotonic logic?

Stenning & van Lambalgen argue that we misrepresent the scope of current logical methods, noting that a range of nonmonotonic logics, in which adding a premise may require withdrawing a previously held conclusion, might meet the challenge of uncertainty. As noted in BR, and elsewhere (e.g., Oaksford & Chater 1991; 2002), there are, however, fundamental problems for nonmonotonic logics in the crucial case where different “lines of argument” clash. Thus, if it is sunny, John goes to the park, and it’s sunny appears to provide a powerful argument that John goes to the park. But adding the premise, John is arrested by the police in a dawn raid, together with background knowledge, appears to yield the conclusion that John does not go to the park.

From the perspective of classical logic, this situation is one of contradiction – and what is needed is a way of resolving which premise should be rejected. For example, one might claim that the conditional if it's sunny, John goes to the park is false, precisely because of the possibility of, among other things, arrest. But, as noted in section R2.1, it is difficult to avoid the conclusion that all conditionals, outside mathematics, are false, because the possibility of counterexamples always exists. Reasoning from premises known to be false is not, of course, justified, whether in logic, or any other standard framework, and hence, the logical analysis of the original argument collapses.

The strategy of nonmonotonic logic attempts to solve this problem by treating the conditional as a default rule, which holds, other things being equal. Indeed, outside mathematics, almost all rules are default rules. Indeed, the implicit rule that allows us to infer that being arrested is incompatible with a trip to the park is itself a default rule, of course – for example, arrest may be extremely brief, or perhaps the police station is itself in the park. Thus, from this viewpoint, uncertain reasoning centrally involves resolving clashes between default rules. In BR, we argue that resolving such clashes is not typically possible by looking only at the structural features of arguments. Instead, it is crucial to differentiate stronger and weaker arguments, and degrees of confidence in the premises of those arguments. Logical methods provide no natural methods for expressing such matters of degree; but dealing with degrees of belief and strength of evidence is the primary business of probability theory.

R3.4. Is logic relevant to cognition?

Several commentators suggest that the powerful machinery of logic should not be jettisoned prematurely (Allott & Uchida, De Neys, O’Brien, Politzer & Bonnefon, Stenning & van Lambalgen). As we noted in section R3.1, probability theory (i.e., modern inductive logic) is a generalization of logic, allowing degrees of uncertainty. However, it is a generalization that is presently limited in scope. This is because how probability interacts with richer representations involving, for example, relations, quantification, possibility, deontic claims, tense and aspect, and so on, is yet to be worked out. BR has, as we have mentioned, some preliminary suggestions about the probabilistic representation of individual connectives (if…then…) and quantifiers (Most, Few, Some, etc.). But this is very far from a full probabilistic generalization of, for example, the predicate calculus, the workhorse of classical logic and natural language semantics. The formal challenges here are substantial. Nonetheless, much progress has been made, in a number of directions, in fusing together probabilistic and logical methods (e.g., see papers in Williamson & Gabbay 2003), thus advancing Carnap’s (1950) program of building an inductive logic. Pfeifer & Kleiter apply logic in an interesting, but distinct, way: as providing a machinery for reasoning about probability, rather than using probability to generalize logic.

According to De Neys, concentrating on the computational level means that BR underplays the role of logic in human reasoning. De Neys argues that latency and brain imaging studies, investigating the mental processing involved in reasoning, rather than just the output of these processes, consistently reveal a role for logic. Yet all the cases that De Neys cites involve a conflict between belief and logic such that prior belief suggests one response, but logical reasoning from the given premises suggests another. However, the Bayesian approach can explain at the computational level why such conflicts might arise and therefore why inhibitory processes might need to be invoked (De Neys et al. 2008; Houdé et al. 2000). Oaksford and Hahn (2007) point out that probabilistic validity of an argument and its inductive strength can conflict. So, for example, Modus Ponens (MP) is probabilistically valid. However, if the probability of the conditional is low, then the inductive strength of the argument, that is, the probability of the conclusion given the premises, will also be low. The right computational level analysis may, therefore, remove the need to propose two special purpose cognitive systems operating according to different principles. This view seems consistent with the current state of imaging studies, which provide little evidence for a dedicated logical reasoning module (Goel 2007).

O’Brien describes Chrysippus’ dog’s ability to follow a scent down one path in a fork in the road, having eliminated the other as an application of the logical law of disjunction elimination – and hence, suggests that logic is cognitively ubiquitous. However, this logical law cannot uncritically be imported into a theory of canine cognition. For one thing, such patterns of behavior are at least as well modeled in probabilistic (Toussaint et al. 2006), as in logical, terms. Indeed, probabilistic methods are crucial in planning tasks in uncertain environments, which is, of course, the normal case, outside mathematically specified game-playing environments. In any case, just because a behavior can be described in logical or probabilistic terms does not directly imply that it is governed by logical or probabilistic processes. The issues here are complex (see the excellent introductory chapter to Hurley & Nudds 2006) and many possibilities would
need to be ruled out before abandoning Lloyd Morgan’s canon: that lower-level explanations of animal behavior should be preferred.

In short, we believe that cognitive science ignores logic at its peril – logic provides powerful and much needed tools, just as do other branches of mathematics. It does not, however, readily capture patterns of human reasoning, or, we suggest, cognition at large, unless generalized into a probabilistic form able directly to deal with uncertainty.

**R3.5. Why probability rather than other numerical measures?**

Danks & Eberhardt and Politzer & Bonnefon ask why we use probability, rather than other numerical measures of degrees of belief, such as confidence intervals, Dempster-Shafer belief functions (Dempster 1968; Shafer 1976), or fuzzy logic (Zadeh 1975). In BR, our primary motivation is practical: Bayesian probabilistic methods provide a natural way to capture human reasoning data; and more generally, Bayesian methods have swept through the brain and cognitive sciences, from understanding neural coding (Doya et al. 2007), through vision, motor control, learning, language processing, and categorization. Even within research on reasoning, Bayesian methods have proved central to understanding inductive inference (Griffiths & Tenenbaum 2005; Tenenbaum et al. 2007), causal reasoning (Sloman 2005; Tenenbaum & Griffiths 2001), and argumentation (e.g., Hahn & Oaksford 2007), as well as the primarily deductive reasoning problems considered in BR. Moreover, probabilistic methods connect with rich literatures concerning computational inference methods (e.g., based on graphical models, Lauritzen & Spiegelhalter 1988; Pearl 1988), machine learning (e.g., Jacobs et al. 1991), and normative theories of reasoning about causality (Pearl 2000). Finally, probability also has deep relationships to other powerful concepts in the brain and cognitive sciences, including information theory (e.g., Blakemore et al. 1991) and simplicity, for example, as captured by Kolmogorov complexity theory (e.g., Chater 1996; Chater & Vitányi 2002). Thus, our focus on probability is primarily pragmatic rather than, for example, depending on a priori justifications.

Danks & Eberhardt focus, nonetheless, on justification, arguing that doubt can be cast on justifications such as the Dutch Book argument and long run convergence theorems. We see the project of rational analysis as a user of probability, on a par with the rest of science, for example, statistical mechanics, Bayesian image restoration, or economics. We only need to be as concerned about justification as these other endeavors. Danks & Eberhardt’s worries are analogous to Berkeley’s objections to Newton’s infinitesimals: of considerable conceptual importance, but with little direct impact on the practical conduct of science. Nonetheless, probability is at least better justified than alternative formalisms for modeling uncertainty.

Politzer & Bonnefon and Danks & Eberhardt raise the possibility that the assumptions of the probabilistic approach may be too strong. We instead believe that they are, if anything, too weak; that is, they define minimal coherence conditions on beliefs, which need to be supplemented with richer formalisms, including, as noted in section R3.4, the ability to represent relations and quantification, and to represent and manipulate causal relations (e.g., Pearl 2000).

**R3.6. Are we Bayesian enough?**

Other commentators (Over & Hajchristidis, Pfeifer & Kleiter, Stenning & van Lambalgen) have the opposite concern: that BR is not Bayesian enough. Over & Hajchristidis argue that in conditional inference, not only is the conditional premise (e.g., if p then q) uncertain, but so is the categorical premise, p. In BR (p. 121), we mention this general case (implying Jeffrey’s rule [Jeffrey 1983]), but point out that this extra element of uncertainty appears unnecessary to capture the conditional reasoning data.

Stenning & van Lambalgen and Pfeifer & Kleiter also argue, in different ways, that we are insufficiently Bayesian. Stenning & van Lambalgen argue that our account of suppression effects is not Bayesian because coherent Bayesian revision of the probability space assumes “rigidity”: that is, the conditional probability P(q|p) remains the same if we learn the truth of a categorical premise: p, q, not-p, or not-q (and no other information). We agree. But this does not imply that P(q|p) remains the same if we are told about that p, because pragmatic factors allow us to infer a great deal of additional information; and this information can legitimately change P(q|p). It is this latter case that is relevant for reasoning with verbal materials. Thus, suppose I believe if the key is turned, the car starts; and I am told: “the car didn’t start this morning.” This would be a pragmatically pointless remark, if the key had not been turned. I therefore infer that the key was turned, and the car didn’t start for some other reason. Thus, I revise down the probability of the relevant conditional P(car starts|key turned) dramatically. So the violation of rigidity, notably in this type of Modus Tollens (MT) inference, does not violate Bayesian precepts, but merely applies them to the pragmatics of utterances (see BR, pp. 126–128; Sobel 2004; Sober 2002).

Pfeifer & Kleiter suggest that inference can proceed locally and deductively in a mental probability logic. In such a logic, the precise probability of a conclusion cannot typically be deduced from the probabilities of the premises – but a probability interval can be. We adopted a similar approach to probabilistic validity for syllogisms where, according to our probabilistic semantics, quantifiers describe probability intervals. Nonetheless, in line with Stanovich and West’s (2000) “fundamental computational bias,” we believe that people spontaneously contextualize and elaborate verbal input, by adding information from world knowledge. Indeed, it takes substantial cognitive effort not to do this. Consequently, we think it unlikely that people reason deductively about probability intervals.

**R3.7. Measuring confirmation**

People are not merely passive observers. They can actively search for information to help test hypotheses, or to
achieve specific goals. In BR, we outline “rational” accounts for both cases. Where people test between hypotheses, a natural objective is to search for data in order to maximize the expected amount of information that will be gained in the task (Shannon & Weaver 1949). This is “disinterested” inquiry. Where people gain information to help achieve specific goals, then a natural objective is to choose information to maximize expected utility (balancing costs of information search with the improved choices that may result from new information). This is “goal-directed” inquiry. In BR, we note that different variations of Wason’s selection task are appropriately captured by versions of one or other model. In particular, we showed how attending to the goal-directed case avoids the postulation of specific machinery, such as “cheater-detection” modules (e.g., Cosmides 1989), to explain patterns of experimental data (e.g., BR, pp. 191–98).

Focusing on disinterested inquiry, Nelson notes that a wide range of normative and descriptive proposals for assessing the strength of information in a piece of data have been proposed. In testing these models against a wide range of psychological data (Nelson 2005), he finds that the information-theoretic measure implicit in our analysis stands up well against competitors, although it is not picked out uniquely by the empirical data.

Poletiek notes a further interesting link to philosophy of science, noting that Popper’s measure of severity of test is equivalent to $P(e|H)/P(e)$, for data $e$ and hypothesis $H$. And the logarithm of this quantity just is the amount of information carried by the evidence $e$ about $H$ — the quantity which we use in our model of disinterested inquiry in the selection task. This quantity is also used as a measure of the degree to which a theory is confirmed by the data in confirmation theory (Milne 1996). This is, as Poletiek notes, particularly interesting, given that Popper’s measure of severity of test is part of a theoretical framework which aims to entirely avoid the notion of confirmation (see also Milne 1995). Thus, our account of the selection task could be recast, from a Popperian standpoint, as a rational analysis in which people attempt to choose data to provide the more severe possible tests for their hypotheses.

**R4. Rational analysis, algorithmic processes, and neural implementation**

BR is primarily concerned with the rational analysis of human reasoning (e.g., Anderson 1990; 1991a; Chater & Oaksford 2008a; Oaksford & Chater 1998b). In this section, we consider the role of rational analysis in the brain and cognitive science and whether this style of explanation is fundamentally flawed.

**R4.1. The power of rational analysis**

Hahn notes that the shift away from considerations of algorithms and representations, encouraged by rational analysis, has led to a substantial increase in explanatory power in cognitive science, in a number of domains. Where the underlying explanation for an aspect of cognition arises from the rational structure of the problem being solved, there focusing on specific algorithmic and neural mechanisms may be unhelpful. Therefore, building specific algorithmic models (e.g., connectionist networks) of a phenomenon may replicate the phenomenon of interest (by virtue of being an adaptive solution to the “rational” problem in hand), but may throw little light on why it occurs.

**R4.2. Normativity and rational analysis**

Evans and Schroyens are concerned about the normative aspect of rational analysis. Evans questions whether normativity is a proper part of a computational-level analysis of human reasoning, and by implication, cognition in general, and recommends a switch to an ecological notion of rationality. He suggests rationality should concern how well people are adapted to their environment, which may not require following the prescriptions of any normative theory of reasoning (cf. Gigerenzer & Goldstein 1996).

We suggest, however, that ecological rationality does not replace, but rather, complements normative rationality. Normative considerations are still required to explain why a particular algorithm works, given a particular environment; indeed, this is precisely the objective of rational analysis. Thus, for example, in arguing for the ecological rationality of various fast and frugal heuristics (Gigerenzer et al. 1999), Gigerenzer and colleagues appeal to a Bayesian analyses to explore the type of environmental structure for which their algorithms succeed (e.g., Martignon & Blackmond-Laskey 1999). Thus, rational analysis cannot be replaced by, but seeks to explain, ecological rationality.

Note, too, that rational analysis is goal-relative: it specifies how best to achieve a given goal, in a given environment, with given constraints (Anderson 1990; Oaksford & Chater 1998b). So, if your goal is to land a rocket on the moon, your guidance system ought to respect classical physics; if your goal is to avoid contradictions, you ought to reason according to standard logic; and if your goal is to avoid accepting bets that you are bound to lose, you ought to follow the rules of probability theory (see Ch. 2 of BR).

Ignoring the goal-relativity of rational analysis leads Schroyens to suggest that we have fallen into Moore’s (1903) “naturalistic fallacy” in ethics: that we have attempted to derive an “ought” from an “is.” Moore’s concern is that no facts about human behavior, or the world, can justify an ethical theory. Ethics is concerned with non-relative notions of “ought”; the aim is to establish universal principles of right behavior. But the goal-relativity of rational analysis makes it very different from the domain of ethics, because it is conditional. Rational analysis considers: if you have objective O, given an environment E, and constraints C, then the optimal action is A. Ethics, by contrast, considers whether O is a justifiable objective. And the nature of the solution to a well-specified optimization problem is itself firmly in the domain of facts.

Indeed, were Schroyens’ concern valid, then its consequences would be alarming, sweeping away functional explanation in biology and rational choice explanation in economics. Yet in all cases, rational/optimality explanations are used to derive empirical predictions; and, as in any scientific enterprise, the assumptions of the rational/optimality accounts are adjusted, where
appropriate, to give a better fit with empirical predictions. Specifically, empirical data lead to revision of empirical assumptions in the rationality/optimality analysis – the empirical data does not lead to a revision of the laws of logic, probability, or any other rational theory.

Khalil raises the opposite concern: that we use rational explanation too narrowly. He argues that the style of optimality explanation that we advocate applies just as well in the explanation of non-cognitive biological structures as it does to cognitive processes – he argues that, in the sense of rationality used in BR, stomachs are just as rational as cognitive mechanisms. This concern appears purely terminological; we reserve “rationality” for information processing systems. But rational analysis is, indeed, parallel to optimality explanation in biology (Chater et al. 2003).

R4.3. Relevance of the algorithmic level

McKenzie and Griffiths note, however, that advocating rational analysis does not make the challenges concerning algorithmic, and indeed neural, implementation disappear. Moreover, the mapping between levels of explanation need not necessarily be straightforward, so that a successful probabilistic rational analysis of a cognitive task does not necessarily require that the cognitive system be carrying out probabilistic calculations – any more than the bird is carrying out aerodynamic calculations – any more than the bird is carrying out aerodynamic calculations – any more than the bird is carrying out aerodynamic calculations – any more than the bird is carrying out aerodynamic calculations – any more than the bird is carrying out aerodynamic calculations – any more than the bird is carrying out aerodynamic calculations. Nonetheless, in many contexts, it is natural to see cognition as carrying out probabilistic calculations; and a priori rational analysis (or, in Marr’s [1982] terms, computational level of explanation) is extremely valuable in clarifying what calculations need to be carried out. Without a “rational analysis” for arithmetic calculations (i.e., a mathematical theory of elementary arithmetic), understanding which algorithms might be used by a pocket calculator, let alone how those algorithms might be implemented in silicon, would be impossible. Griffiths outlines key challenges for creating an algorithmic-level theory of cognition, viewed from a Bayesian perspective; and this perspective dovetails nicely with work viewing neural machinery as carrying out Bayesian inference (e.g., Ma et al. 2006; Rao et al. 2002), which we consider briefly further on.

BR is largely focused on rational level explanation (Anderson 1990; 1991a). Indeed, following Marr (1982), we argued that understanding the rational solution to problems faced by the cognitive system crucially assists with explanation in terms of representations and algorithms, as stressed by Hahn and Griffiths. In BR, this is illustrated by our model of syllogistic reasoning, which proposes a set of “fast and frugal” heuristics (Gigerenzer & Goldstein 1996) for generating plausible conclusions, rooted in a Bayesian rational analysis (Chater & Oaksford 1999b). More recently, we have suggested methods for causal and conditional reasoning, based on “mental mechanisms” (Chater & Oaksford 2006; Ali et al., in press) directly building on rational and algorithmic models inspired by the literature on Bayesian networks (Glymour 2001; Pearl 1988; 2000). Moreover, an explicit algorithmic implementation of our probabilistic account of conditional inference has been constructed using a constraint satisfaction neural network (Oaksford & Chater, in press). Moreover, there is a significant movement in current cognitive science that focuses on developing and employing Bayesian machine learning techniques to model cognition at both the rational and algorithmic levels (e.g., Griffiths et al. 2007; Kemp & Tenenbaum 2008).

Evans’ concern that we ignore the algorithmic level is therefore puzzling. He worries that BR recommends that one should “observe some behaviour, assume that it is rational, find a normative theory that deems it to be so, and then ... nothing else, apparently.” We assume that the ellipsis should, in Evans’ view, be fleshed out with an algorithmic, process-based explanation, which should then be subject to rigorous empirical test. The abovementioned list of algorithmic level proposals inspired by Bayesian rational analysis, both in the domain of reasoning and in cognitive science more generally, gives grounds for reassurance. Moreover, the extensive empirical testing of these models (Green & Over 1997; 2000; McKenzie & Mikkelsen 2000; 2007; McKenzie et al. 2001; Nelson 2005; Oaksford & Moussakowski 2004; Oaksford & Wakefield 2003; Oaksford et al. 1999; 2000; Tenenbaum 1999) should allay concerns that rational analysis provides no testable predictions. Ironically, the only theories in the psychology of reasoning that have been algorithmically specified, aside from those within the Bayesian tradition, are directly based on another rational level theory: logic (Johnson-Laird 1992; Rips 1994). Theorists who have instead focused primarily on heuristics for reasoning have couched their explanations in purely verbal terms (Evans 1989; Evans & Over 2004). This indicates, we suggest, that rational analysis assists, rather than impedes, algorithmic explanation.

R4.4. Relevance of neural implementation

Bayesian rational analysis is, moreover, appealing because it appears to yield algorithms that can be implemented in the brain. In BR (Ch. 4), we observed that the Bayesian approach was sweeping across cognitive psychology. We might also have added that its influence in computational neuroscience is at least as significant (Friston 2005). Although our Bayesian analyses of higher-level reasoning do not directly imply Bayesian implementations at the algorithmic level, it is intriguing that influential theorists (Doya et al. 2007; Friston 2005; Ma et al. 2006) view Bayesian inference as providing the driving computational principle for neural information processing. Such models, using population codes (Ma et al. 2006), which avoid treating the brain as representing probabilities directly on a numerical scale, can model simple perceptual decision tasks (Gold & Shadlen 2000). Such convergence raises the possibility that Bayesian rational analyses of reasoning may one day find rather direct neural implementations.

De Neys specifically appeals to the implementation level in commenting on BR. He draws attention to imaging studies of reasoning that suggest a role for the anterior cingulate cortex in detecting conflict and inhibiting responses. As we have seen (sect. R3.4), such a role is entirely consistent with Bayesian approaches. Indeed, more broadly, imaging work on human reasoning, pioneered by Goel (e.g., Goel 2007), is at an exploratory stage, and currently provides few constraints on theory. Moreover, as we have seen, where cognitive
neuroscientists concentrate on what computations the brain performs rather than where, the emerging answer is Bayesian.

**R4.5. Optimality and rational analysis**

A range of commentators (e.g., Brighton & Olsson, Danks & Eberhardt, Evans, and Schroyens) argue that the methodology of rational analysis faces conceptual problems. Our general response to these concerns is pragmatic. As with any methodology, we see rational analysis, using probabilistic methods or otherwise, as primarily to be judged by its results. Anderson’s path-breaking work (1990; 1991a), and the huge literature on Bayesian models across the brain and cognitive sciences, of which *BR* is a part, is therefore, in our view, the best argument for the value of the approach. Parallels with closely related work in behavioral ecology and rational choice explanation in economics give further weight to the view that a “rational” style of explanation can yield considerable insights. But, like any style of explanation, rational analysis has its limits. Just as, in biology, some behaviors or structures are products of “history” rather than adaptation (Carroll 2005), and some economic behaviors are the product of cognitive limitations (e.g., Ariely et al. 2003; Thaler 2005), so in the brain and cognitive sciences, we should expect some phenomena to arise from specific aspects of algorithms/representations or neural implementation.

We are therefore happy to agree with commentators who suggest that there are cognitive phenomena for which purely rational considerations provide an incomplete, or indeed incorrect, explanation (e.g., Brighton & Olsson, Evans). We also agree that rational analysis is challenged where there are many, perhaps very different, near-optimal rational solutions (Brighton & Olsson). In such situations, rational analysis provides, at best, a range of options — but it does not provide an explanation of why one has been chosen. Nonetheless, these issues often cause few problems in practice, as the results in *BR* and in the wider program of rational explanation illustrate.

We agree, moreover, with concerns that finding exactly the optimal solution may be over-restrictive (Brighton & Olsson, Evans). Consider the case of perceptual organization, where the cognitive system must decide between multiple interpretations of a stimulus (Gregory 1970; von Helmholtz 1910/1925). Accounts based on Bayesian probability and on the closely related idea of maximizing simplicity (Chater 1996; Hochberg & McAlister 1953; Leeuwenberg & Boselie 1988) adopt the perspective of rational analysis, but they do so comparatively. That is, the perceptual system is presumed to choose interpretation A, rather than interpretation B, if A is more likely than B (or, in simplicity-based formulations, if it provides a simpler encoding of the sensory input). Neither the likelihood nor the simplicity principles in perceptual organization are presumed to imply that the perceptual system can optimize likelihood/simplicity — and indeed, in the general case, this is provably impossible (see Chater 1996, for discussion). Indeed, we suspect that rational analysis will, in many cases, primarily be concerned with providing a measure of the relative “goodness” of different cognitive processes or behaviors; and it is explanatory to the degree to which the “good” mechanisms are more prevalent than the “bad.” The parallel with evolutionary explanation seems to be exact here: Inclusive fitness provides a crucial explanatory measure in explaining the evolution of biological structures, but the explanatory “bite” is comparative (i.e., in a certain environment, a flipper yields greater fitness than a leg). There is no assumption that biological evolution, in any context, reaches a state of completely optimized perfection; indeed, quite the reverse (Jacob 1977). Thus, Evans’ emphasis on satisficing rather than optimizing, and Brighton & Olsson’s focus on relative rationality, seem to us entirely consistent with *BR*.

Note, too, that in modeling many aspects of cognition, a full-scale rational analysis (specifying a task, environment, and computational limitations) may not be required. For example, conditional inference can be modeled in Bayesian terms, assuming only a probabilistic interpretation of the premises, and the requirement of maintaining consistent degrees of belief. The success of the probabilistic, rather than a logical, interpretation of the premises can be assessed by comparing the predictions of both approaches to data on human reasoning, as well general philosophical principles.

*Brighton & Olsson* also raise a different concern: that the specific sets of probabilistic assumptions (such as the independence assumptions embodied in naïve Bayes) may sometimes be justified not by rational analysis, but instead in the light of their general, formal properties, combined with empirical success in solving some externally defined task (e.g., estimating the relative sizes of German cities, Gigerenzer & Goldstein 1996). For example, a model such as naïve Bayes, they note, may be effective because it has few parameters and hence avoids over-fitting. We suggest, however, that this is not a separate type of explanation of inferential success, distinct from Bayesian rational analysis. Instead, the justification for preferring simple models can, itself, be provided in terms of Bayesian reasoning, and closely related formalisms, including minimum description length (Chater & Oaksford 2005b; MacKay 2003; Rissanen 1989; Vitányi & Li 2000).

**R4.6. Need rational explanation be causal?**

*Brighton & Olsson*, together with Danks & Eberhardt, raise the fundamental concern that rational explanation does not provide a causal explanation of behavior. We agree. Rational explanation is teleological (Fodor 1968) — it explains by reference to purpose, rather than cause.

In particular, rational explanation does not require that the rational analysis is itself represented in the mind of the agent, and does not, therefore, imply that behavior is governed by any such representation. Aerodynamics may provide an optimality-based explanation of the shape of the bird’s wing; but aerodynamic calculations by the bird (or any other agent) are not causally responsible for the wing’s shape.

Similarly, delineating the circumstances in which algorithms such as naïve Bayes (*Brighton & Olsson*; Domingos & Pazzani 1997), Take the Best (Gigerenzer & Goldstein 1996; Martignon & Hoffrage 1999), or unit-weighted regression (Dawes 1979) are reliable may require highly sophisticated rational explanation. Yet a
cognitive system that employs such models may know nothing of such rational explanations – and indeed, these rational assumptions typically play no causal role in determining the behavior. Thus, in behavioral ecology, for example, the strategies animals use in foraging, mate selection, and so on, are typically explained using optimality explanations; but animals are not assumed to carry out optimality calculations to validate their behavioral strategies.

Danks & Eberhardt suggest that there is a “requirement for a teleological explanation that the normative principle must have played a causal role – ontogenetic, phylogenetic, or both – in the behavior’s existence or persistence. ‘Origin stories’ are required for teleological explanation.” We find this claim puzzling: normative principles, and rational explanations in general, are abstract – they are not part of the causal realm. Thus, a Bayesian rational analysis can no more cause a particular piece of behavior or reasoning, than the principles of arithmetic cause a calculator to display a particular number. Teleological explanations are distinctively non-causal, and necessarily so.

In this section, we have considered concerns about the general project of rational analysis. We now turn to consider specific issues relating to the rational models and empirical data presented in BR.

R5. Reconsidering models and data

Even if the broad sweep of arguments from the preceding sections is endorsed, there remain doubts about the details of the particular models described in BR and their ability to account for human reasoning data. Indeed, in the commentaries, issues of detail emerge most often between researchers who otherwise are in broad agreement. It is in this light that we consider the comments of Liu, Oberauer, Over & Hadjichristidis, and Wagenmakers. We also consider here Halford’s comments on syllogistic reasoning, drawn from a different framework.

R5.1. Conditional inference

Liu, Oberauer, and Over & Hadjichristidis, who have also advocated a probabilistic approach (in particular, to conditional inference), have concerns about our specific model. We addressed, in section R3.6, Over & Hadjichristidis’s argument that we are not Bayesian enough, and that we should employ Jeffrey’s rule to deal with uncertainty in the categorical premise of conditional inference. We pointed out that we too explicitly adopted Jeffrey’s rule in BR. They also cite some unpublished results apparently showing that people have an imperfect understanding of Jeffrey’s rule. These results are intriguing and suggest that more extensive empirical testing of this rule is required.3

Oberauer argues that our models of conditional inference and data selection may lead to absurdity. He argues that if the marginals, P(p) and P(q), remain fixed, which he describes as “axiomatic” in our theory,4 then if one increases the probability that someone gets a headache, given they take drug X, then those who don’t take X will get fewer headaches. This apparent absurdity stems from a conflation in Oberauer’s description between the factual and the epistemic/doxastic: Changing this conditional degree of belief does not mean that these people actually achieve these benefits. In ignorance of the real conditional probability, but knowing the values of the marginals, I should revise my degree of belief that not taking this drug leads to fewer headaches. Yet this will only be appropriate when the marginals are known – which is clearly inappropriate in Oberauer’s example.

Oberauer also perceives an inconsistency between our adoption of The Equation – P(if p then q) = P(q)p – and our use of a contingency table to represent the conditional hypothesis in data selection. However, by The Equation there is only sufficient information in the premises of a conditional inference to draw MP by Bayesian (or Jeffrey) conditionalization (at least a point value). The remaining inferences can only be drawn on the assumption that people use the marginals to calculate the relevant conditional probabilities, for example, P(¬q|¬p) for Denying the Antecedent (DA). Once P(q|p) and the marginals are fixed, the contingency table is determined. Knowing the meaning of a statement is often equated with knowing the inferences that a statement licenses (Dowty et al. 1981). According to The Equation, the conditional only licenses “probabilized” MP. Probabilistically, to draw further inferences requires more information to be drawn from world knowledge. Hence, there is no inconsistency. Moreover, in the selection task, people are presented with an array of possible evidence types that makes the marginals relevant in the same way as presenting more than just MP in the conditional inference task. The degree of belief that is modified by selecting data is in the conditional and the marginals, which constitute the dependence and independence models. Thus, Oberauer’s concerns can be readily addressed.

Oberauer also suggests that contingency tables are consistent with a probabilistic contrast approach, that is, the measure of the strength of an argument, for example, MP, is P(q|p) − P(q|¬p). It is for this reason that we believe that argument strength may indeed be two-dimensional (Oaksford & Hahn 2007). The conditional probability alone can mean that a good argument leads to no increase in the degree of belief in the conclusion, for example, for MP when P(q|p) = P(q) = 1. The probabilistic contrast (and other measures; see, e.g., Nelson, Poletiek, and Oaksford & Hahn 2007) captures the change in the probability of the conclusion brought about by an argument. Oberauer suggests that there is no evidence for people’s use of the probabilistic contrast. Yet Over et al. (2007) found significant sensitivity to P(q|¬p), consistent with some use of the probabilistic contrast or a related measure of change, and the evidence is currently equivocal.

Oberauer also raises two concerns over evidence for our model of conditional inference. First, fitting a model with two free parameters to four data points “is no convincing accomplishment.” Even so, as Hahn observes, the move to detailed model fitting of quantitative data represents significant progress in the psychology of reasoning (for early examples, see Krauth [1982] and Klauer [1999]). Moreover, in BR (pp. 146–49) we fitted the model to the 32 data points produced in Oaksford et al.’s (2000) Experiment 1 using only nine parameters, collapsing far more degrees of freedom than the model fitting reported in Oberauer (2006). Although Oberauer (2006) found poorer fits for our model than alternative theories,
Oaksford and Chater (2008) found that the revised model presented in BR may provide better fits to Oberauer’s data. Second, Oberauer argues that the most relevant empirical evidence comes from studies where probabilities were directly manipulated, of which he mentions two, Oaksford et al. (2000) and Oberauer et al. (2004). Moreover, he argues that their results are equivocal. However, several other studies have manipulated probabilities in conditional inference and found evidence in line with a probabilistic account (George 1997; Liu 2003; Liu et al. 1996; Stevenson & Over 1995). Oberauer also leaves aside the many studies on data selection showing probabilistic effects (see BR, Ch. 6).

Liu’s arguments about second-order conditionalization point, we think, to an important factor that we have yet to consider in reasoning, that is, the effects of context. Liu has found that people often endorse the conclusion that, for example, Officer flies on being told that Officer is a bird in the absence of the conditional premise (reduced problems). This occurs because they fill in this information from world knowledge. However, Liu also found that endorsements increase when the conditional premise is added (complete problems). In BR, we argued that this occurs because people take the conditional premise as evidence that the conditional probability is higher (an inference that may arise from conversational pragmatics). Liu argues that our account implies that manipulations affecting reduced problems should also affect complete problems and provides evidence against this. Yet context, both cognitive and physical, may explain these differences in a way similar to recent studies of decision-making (Stewart et al. 2006). For example, suppose one is told about two swanneries, both containing the same number of swans. In one, 90% of swans are black; in the other, 90% are white. On being told that is a swan, presumably one would only endorse if .5. This is because conversational pragmatics and world knowledge indicate that is in one of the just mentioned swanneries, but the dialogue up to this point does not indicate which one. However, the addition of the conditional premise if a bird is a swan it is white immediately indicates which swannery is being talked about, that is, the one in which is white. Clearly, although manipulations of the relative number of swans in each swannery might affect the reduced problem, they should not affect the complete problem. So if the swannery in which most swans are black were one tenth of the size of the other swannery, then, given natural sampling assumptions, endorsements for the reduced problem should increase to .83, but endorsements of the complete problem should remain the same.

R5.3. Syllogisms and development

Halford argues that mental models theory and a relational complexity measure fit the data as well as the probability heuristics model (PHM); conceding, however, that only PHM generalizes to most and few. Copeland (2006) has also recently shown that PHM provides better fits than mental models and mental logic for extended syllogisms involving three quantified premises. Halford also suggests that basing confidence in the conclusion on the least probable premise, as in our max-heuristic, is counterintuitive. He proposes that confidence should instead be based on relational complexity, which covaries with the least probable premise. But perhaps Halford’s intuition goes the wrong way: the least probable premise is the most informative; and surely the more information you are given, the stronger the conclusions you can draw?

De Neys and Straubinger, Cokely, & Stevens (Straubinger et al.) both argue that there are important classes of evidence that we do not address. De Neys argues that attention to latency data and imaging studies provides a greater role for logic, a claim we disputed earlier. Note, also, that the algorithmic theory in PHM has been applied to latency data and accounts for the data, as well as mental models (Copeland & Radvansky 2004). Straubinger et al. are concerned that we ignore developmental data. In particular, they view the findings on the development of working memory as providing a particular challenge to a Bayesian approach. They do, however, acknowledge that in different areas (e.g., causal reasoning), Bayesian ideas are being successfully applied to developmental data (Navarro et al. 2006; Sobel et al. 2004). Straubinger et al.’s emphasis on working memory provides good reason to believe that our particular approach to deductive reasoning may extend to development. Copeland and Radvansky (2004) explicitly related working-memory limitations to PHM, finding that it provided as good an explanation as mental models theory of the relationship between working-memory capacity and reasoning performance. This result provides some indication that, at least for syllogistic reasoning, developmental trajectories explicable by mental models may be similarly amenable.

R5.2. Data selection

Wagenmakers raises a variety of concerns about our optimal data selection model. First, why do we concede that people should select the standard “logical” A card and 7 card choices, if the rule only applies to the four cards? In BR (p. 210), we argue that people rarely use conditionals to describe just four objects – they assume that the cards are drawn from a larger population. Consequently, we quite explicitly do not make the counterintuitive prediction that Wagenmakers ascribes to us. Second, Wagenmakers wonders why – when all cards carry some information – do participants not select all the cards, if they are maximizing information gain? We assume that the pragmatics of the task suggests to participants that they should select some cards, but not others (BR, pp. 200–201). Third, Wagenmakers suggests that incentivized individuals with more time might make the logical response. Work on individual differences (e.g., Stanovich & West 2000) is consistent with the view that logical competence is learned, either directly (e.g., studying logic or math) or indirectly (e.g., learning to program or learning conventional, non-Bayesian statistics); such logical competence is a prerequisite for “logical” responses, and covaries with IQ as measured in University populations. Wagenmakers also remarks that, as Bayesians, we should avoid null hypothesis testing in statistically assessing our models. This choice is purely pragmatic: it conforms to the current demands of most journals.
to explanation in terms of probability heuristics. Our approach also provides a natural way in which experience, leading to the learning of environmental statistics, might influence reasoning development. Exploring these possibilities must await future research.

R6. The Bayesian turn

BR is part of a larger movement across the brain and cognitive sciences – a movement which sees cognition as centrally concerned with uncertainty; and views Bayesian probability as the appropriate machinery for dealing with uncertainty. Probabilistic ideas have become central to theories of elementary neural function (Doya et al. 2007), motor control (Kording & Wolpert 2004), perception (Knill & Richards 1996), language processing (Manning & Schütze 1999), and high-level cognition (Chater & Oaksford 2005a; Chater et al. 2006). They also cut across Marr’s (1982) computational (Anderson 1990; Pearl 2000), algorithmic (Jacobs et al. 1991), and implementational (Doya et al. 2007) levels of explanation. In arguing that commonsense reasoning should be understood in terms of probability, we are merely recasting Laplace’s (1814/1951) classic dictum concerning the nature of probability theory: “The theory of probabilities is at bottom nothing but common sense reduced to calculus.”

NOTES

1. Although Braine and O’Brien (1991) explicitly reject the use of relevance logic (Anderson & Belnap 1975), this does provide an interesting possible route for developing these ideas. In particular, interpretations of the semantics of relevance logics as a ternary relation between possible worlds, or from an information-theoretic perspective, as a ternary relation between a source, a receiver, and a channel (Restall 1996), may provide interesting connections with nonmonotonic reasoning.

2. By contrast, we know of just one paper in the psychology of reasoning discussing Dempster-Shafer belief functions, namely, George (1997).

3. Its normative status has also been questioned for many years (see, e.g., Field 1978).

4. This is despite the fact that they were not fixed in Oaksford and Chater (1994).

5. Of course, different assumptions would yield different results. For example, if the previous dialogue had been talking about the swannery, where most swans are black, just before introducing Tweety, the assumption may be that Tweety comes from that swannery and so *Tweety is white* might only be endorsed at .1.

References

[The letters “a” and “e” before author’s initials stand for target article and response references, respectively]


(1998) A primer of probability logic. CSLI Publications. [aMO]


References/Oaksford & Chater: Précis of Bayesian Rationality


Oberauer, K. & Wilhelm, O. (2003) The meaning(s) of conditionals:

Rationality in an uncertain world. Psychology Press. [aMO, WS, ed.]


Pearl, J. (1988) Probabilistic reasoning in intelligent systems. Morgan Kaufmann. [aMO]


