

# A Normative Theory of Argument Strength

ULRIKE HAHN

*Cardiff University*

MIKE OAKSFORD

*Birkbeck College London*

**Abstract:** In this article, we argue for the general importance of normative theories of argument strength. We also provide some evidence based on our recent work on the fallacies as to why Bayesian probability might, in fact, be able to supply such an account. In the remainder of the article we discuss the general characteristics that make a specifically Bayesian approach desirable, and critically evaluate putative flaws of Bayesian probability that have been raised in the argumentation literature.

**Résumé:** Dans cet article nous soutenons l'importance générale des théories normatives de l'appui des raisons d'un argument. Nous employons nos recherches récentes sur les sophismes pour appuyer l'idée que la probabilité bayésienne pourrait contribuer à ces théories. Ensuite nous discutons des caractéristiques générales qui suggèrent les avantages d'une approche bayésienne, et évaluons les présumés points faibles attribués aux probabilités bayésiennes et avancés dans les publications sur l'argumentation.

**Keywords:** argumentation, fallacies, Bayesian probability, argument from ignorance, slippery slope arguments, circular arguments, Theophrastus rule

## Introduction

Some arguments we find convincing, others we do not. Is such an evaluation simply a matter of arbitrary preference or can rational justifications for such choices be given? Can there be such a thing as a normative theory of argument strength? It is the contention of this paper that such a theory is not only highly desirable, but also possible. Specifically, our recent work on the fallacies (Oaksford & Hahn, 2004; Hahn & Oaksford, in press; Hahn, Oaksford & Bayindir, 2005; Hahn, Oaksford & Corner, 2005; and see also, independently, Korb, 2004), has sought to develop a general account of the fallacies based on Bayesian probability. The results of this, we will argue, give some hope that Bayesian probability might be able to provide a general, normative theory of argument strength. In how far this will be possible, of course, currently remains unclear. The question can only fully be addressed by actual demonstration of the account's adequacy in a wide range of circumstances; consequently there is a wealth of further research to be done. Here, we seek only to provide an argument for why investment in this project seems worthwhile. This argument has four main parts. In the first part of the paper we outline general considerations as to why a normative account of argument

strength would be desirable. In the second part, we give a brief overview of our results to date, in order to establish the case that a Bayesian theory looks at least in contention. In the third, and main, section we draw out the particular properties of Bayesian probability that we think make a specifically Bayesian account of argument strength attractive. Finally, some arguments against a Bayesian approach that have been voiced in the argumentation literature are addressed, and it is argued that the seeming limitations of Bayesian probability are in fact strengths.

### **1. Why a normative theory of argument strength?**

Two obvious answers dominate here: first, such a theory is of inherent theoretical interest and, second, it is of obvious applied importance. Interest in standards of rational inference and hence argument has, in one way or another, motivated research on logic for much of its history (see Prakken & Vreeswijk, 2002 for an overview of recent work concerned with natural language argumentation). Within the last decades, numerous authors have come to doubt that logic could provide an appropriate standard against by which to judge argument strength (e.g., Hamblin, 1970; Heysse, 1997; Johnson, 2000; also Boger, 2005 for further references). Logic's perceived failures have fuelled the rise of dialectical and rhetorical theories (see e.g., Slob, 2002 for discussion). These theories have focussed on properties of discourse, not the evaluation of the inherent qualities of sets of reasons and claims. Nevertheless, proponents of such theories have frequently sought to use discourse rules to evaluate at least some classes of arguments, in particular classic fallacies, as good or bad (e.g., Walton, 1995; van Eemeren & Grootendorst, 1992, 2004). In this sense, a preoccupation with argument strength has remained even here. On the theoretical side then, whether or not an adequate theory of argument strength is possible, and what it would look like, is a longstanding theoretical question demanding resolution.

At the same time, the applied importance of a theory of argument strength is obvious. In complex societies such as ours argumentation plays a central role. This has led not only to a burgeoning literature concerned with 'critical thinking' and its teaching (e.g., McPeck, 1981; Siegel, 1988; Bowell & Kemp, 2002; Woods, Irvine & Walton, 2004), it has also motivated the development of domain specific theories of argumentation, for example in law (for overviews see, Neumann, 1986; Feteris, 1997). Argumentation in applied settings would be directly affected by the development of a suitable normative theory of argument strength and the capacity for rational resolution it would provide.

Somewhat less obviously, the question of a normative theory of argument strength is relevant also to those pursuing alternative, seemingly opposed paths in the form of consensus theories (e.g., Alexy, 1989; van Eemeren & Grootendorst, 2004 for examples). For one, relative to "right or wrong", "consensus" is typically only second best. The reason researchers default to consensus theories, whether in the domain of theories of truth or in the domain of argumentation, is because the ultimate prize, a normative theory of content, seems unattainable. Were a normative

theory of content available, it is that theory we would look to for conflict resolution.

In fact, the relationship between a normative content theory of argument strength and consensus theories of argumentation is more complex. Inevitably given their focus on a non-content based outcome characteristic –consensus- these theories tend to have a strong emphasis on *procedure*. The rules and norms they posit are rules of engagement: proponents can only put forward claims they actually believe (e.g., Alexy, 1989), proponents must justify claims when challenged (van Eemeren & Grootendorst, 2004) and so on. Crucially, the need for procedural rules remains even where objective standards of content evaluation exist. Even where the goal becomes ‘truth’, ‘the best perspective’ or the ‘strongest position’ there will still be rules of engagement that will make that outcome more or less likely to occur (see also Goldman, 1994). Silencing opponents by force, for example, is undesirable not just with regards to consensus, but also because the suppression of arguments in discourse means that the potentially strongest argument might not be heard. This means the insights of researchers developing consensus theories are unlikely to become obsolete. Some reorientation and adjustment to specific rules of discourse would likely be necessary were a normative theory of content to supplant the emphasis on consensus, but normative theories of content and procedural theories ultimately pursue complimentary goals (Goldman, 1994; Hahn & Oaksford, in press), both of which have an important role to play.

Finally, a normative theory would provide an organizing framework for *descriptive* work. This may seem counterintuitive at first, but attempts at scientific description of human argumentative behavior within cognitive science, and particularly within cognitive psychology would benefit hugely from a normative theory around which to structure research.

Normative theories of behavior and thought provide standards against which actual human performance can be compared. For one, this provides a ready set of questions for descriptive research to address- specifically, how far do human beings match up to these standards and where specifically do they fall short. However, normative standards also play a vital role in interpreting and understanding human behavior. For a cognitive scientist, complete explanation encompasses several, mutually informative levels of description (Marr, 1982). Human cognition as a computational process is to be understood at the hardware level that governs how a process is actually implemented, at a representational or algorithmic level that characterizes the procedures involved, and at the highest level, the so-called computational level, through a general characterization of the problem the cognitive process is seeking to address. Though these levels constrain one another, they also exhibit some degree of independence.<sup>1</sup> Because of this degree of independence, the task of explaining any computational system is not complete before all three levels have been addressed. Normative theories have an important role to play in computational level explanation, particularly so within the so-called *rational analysis* of behavior (e.g., Anderson, 1991; Chater & Oaksford, 2000; Oaksford & Chater, 1998a; Chater & Oaksford; 1999a). Rational analysis seeks to understand human behavior as an approximation to some ideal behavior, typically as an adaptation to

some environmentally posed problem. Rational analysis seeks to characterize a problem faced by the cognitive system, develop what would be the 'best' solution, and then to determine the extent to which system behavior can be seen to be an approximation of that solution, even though it might fall short in certain circumstances. Consequently, normative theories can form part of functional 'why' questions in the analysis of behavior which are essential to the understanding of purposeful behavior.

These considerations are exemplified by the enormous success of the psychological literatures on "naïve statistics", decision-making and logical reasoning (see e.g., Kahneman, Slovic & Tversky, 1982; Holyoak & Morrison, 2005). In all three cases, clear normative theories have prompted obvious research questions with regards to the extent to which human beings conform to normative considerations. This has not only led to considerable insight into what humans can (and cannot) readily do and informed our understanding of human rationality; by provoking explanation of seeming deviations from normative theories it has also led to theories about mechanism and process, that is, theories about the specific cognitive means by which reasoning in these contexts is achieved. Though these areas are not complete, they are, by the standards of psychology, fairly "mature" and developed areas of research. It does not seem unreasonable to suppose that the lack of a clear normative theory of argument strength is one of the reasons why the wider psychological study of argumentation is underdeveloped by comparison.

Experimental research in this area is currently fragmented across a variety of individual specialist domains. There is, on the one hand a wide literature on 'persuasion' which has, by and large, examined non-strength related circumstances in their influence -typically specifically on attitudes (see e.g., Maio & Haddock, in press, or Johnson, Maio & Smith-McLallen, in press, for an overview). There is also, in addition to the logical reasoning literature, a much more narrow literature examining several kinds of inductive arguments (e.g., Osherson et al, 1990). Finally, there are a few studies that have been informed by a broadly pragma-dialectic perspective (Neuman, 2003; Neuman & Weizman, 2003; Neuman, Weinstock & Glasner, in press; Weinstock, Neuman & Tabak, 2004; Rips, 1998, 2002) as well as a body of developmental research (for a review see Felton & Kuhn, 2001).

All of these have proceeded in virtually complete isolation from one another and there is nothing even remotely like an integrated account to be had (but see Rips, 2001 and Oaksford & Hahn, in press, for integration of at least some of these). A normative account of strength could serve to link these different bodies of research, as well as generate novel testable predictions of its own.

There is reason to believe that a specifically probabilistic, normative account would be particularly useful here. In parallel to our development of a normative theory, we have begun a program of experimental research involving the fallacies from a Bayesian perspective (Oaksford & Hahn, 2004; Hahn, Oaksford & Bayindir,

2005), and this work has natural connections to the literature on belief and attitude change under the header of the ‘subjective probability model’ which have yet to be explored (e.g., Allen & Kellermann, 1988; Allen, Burrell & Egan, 2000; Hample, 1977, 1978, 1979; McGuire, 1960; Wyer, 1970, 1974; Wyer & Goldberg, 1970). We detail the (normative) Bayesian reconstruction of fallacies in the next section in order to motivate further the desirability of a Bayesian account.

## **2. Why might Bayes provide a theory of argument strength?**

Support for the idea that Bayesian probability could provide a normative theory of argument strength comes from the fact that it has successfully been used to explain a considerable range of fallacies, in that it captures key intuitions about the relative strength of arguments. In conjunction with Bayesian probability’s pedigree as a normative framework, this suggests it might be able to supply the long-desired formal treatment of the fallacies (Hamblin, 1970). This work can be seen as providing a formal explication of the epistemic account of the fallacies (e.g., Siegel & Biro; 1997; Ikenobe, 2004). Because this work is published elsewhere, we provide only a brief overview here.

Oaksford and Hahn (2004) first sought to explain through a Bayesian account arguments from ignorance such as

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

Individual arguments are composed of a conclusion and evidence for that conclusion. Both conclusion and evidence have associated probabilities which are viewed as expressions of subjective degrees of belief. Bayes’ theorem provides an update rule for the degree of belief associated with the conclusion in light of the evidence. Argument 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.

A Bayesian account captures, among other things, the difference between positive and negative evidence and allows one to capture the intuition that the positive argument (2a) is stronger than the negative argument (2b):

- (2a) *Drug A is toxic because a toxic effect was observed (positive argument).*

- (2b) Drug A is not toxic because no toxic effects were observed (negative argument, i.e., the argument from ignorance).

However, (2b) too can be acceptable where a legitimate test has been performed, i.e.,

If drug A were toxic, it would produce toxic effects in legitimate test.

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  $\neg e$  stand 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:

$$P(T | e) = \frac{P(e | T)P(T)}{P(e | T)P(T) + P(e | \bar{T})(1 - P(T))} \quad (3)$$

$$P(\bar{T} | \bar{e}) = \frac{P(\bar{e} | \bar{T})(1 - P(T))}{P(\bar{e} | \bar{T})(1 - P(T)) + P(\bar{e} | T)P(T)} \quad (4)$$

As Oaksford and Hahn (2004) argue, sensitivity and selectivity, for a wide variety of clinical and psychological tests are such that positive arguments are stronger than negative arguments. The reason we consider negative evidence on ghosts (1) to be so weak is because of the lack of sensitivity (ability to detect ghosts) we attribute to our tests as well as our low prior belief in their existence (see also Hahn et al, 2005 and Hahn and Oaksford, in press, for further discussion and analysis of different kinds of arguments from ignorance). The Bayesian account renders this textbook example as an argument that occupies the extreme lower end of the argument strength range as a consequence of the specific probabilities estimates involved. The argument is weak because of these aspects of its content, not because of its logical structure or particular role in a discourse (cf., Walton, 1996) and other arguments with the same structure and discourse function can be perfectly convincing. In other words, the Bayesian analysis tackles what has been a longstanding problem for the fallacies, namely that most types of fallacy seem prone to a proliferation of exceptions that seem more or less acceptable. The Bayesian account allows one to distinguish 'good' arguments from ignorance from less compelling ones, providing an explanation for why they are good or bad.

Hahn and Oaksford (in press) also provide a Bayesian treatment of slippery slope arguments which as consequentialist arguments are captured using decision theory (Savage, 1954; on decision theory and consequentialist argument see also e.g., Lumer, 1997). According to Bayesian decision theory, choosing an action in the face of an uncertain future should be based on an evaluation both of the utilities

we assign to possible future outcomes and the probabilities with which we think these outcomes will obtain. This normative framework can be applied to textbook slippery slope arguments such as

- (5) We should not ban private possession of automatic weapons, because doing so will be the first step on the way to a communist state.

This argument seems weak, because the subjective probability of such a ban indeed setting in motion a chain of events that will lead to the outcome ‘communist state’ seems so incredibly low. Yet at the same time, there is evidence from legal history that ‘slippery slopes’ have, in fact occurred (Lode, 1999), and, in everyday life, the mechanisms behind them are exploited through techniques such as ‘foot in the door’ advertising (e.g., Freedman & Fraser, 1966). In general, the more there is a real chance of a feared outcome occurring the stronger a slippery slope argument will be. Hence, in

- (6a) Legalizing cocaine will lead to an increase in heroin consumption.

- (6b) Legalizing cannabis will lead to an increase in heroin consumption.

the first argument (a) seems more compelling. That the degree to which one cares about the outcome, its utility, also plays a role can be seen from the examples

- (6c) Legalizing cannabis will lead to an increase in heroin consumption.

- (6d) Legalizing cannabis will lead to an increase in listening to reggae music.

where, assuming that both listening to reggae music and heroin consumption are equally likely, (c) seems the far stronger argument. By varying both utility and probability, perfectly acceptable examples of slippery slope arguments can readily be generated (see for examples also Corner, Hahn & Oaksford, 2006).

Hahn and Oaksford (in press) also provide a Bayesian analysis of the *argumentum ad populum* or “appeal to popular opinion” (on this see also, Korb, 2004) and the *argumentum ad misericordiam*, which uses an appeal to pity or sympathy for argumentative support. In Hahn, Oaksford and Bayindir (2005) the Bayesian account is extended to a treatment of circular arguments, and we say more on these below.

This treatment of informal argument fallacies complements earlier research by Oaksford and Chater that has argued in detail that a wide range of seeming ‘logical errors’ in conditional and syllogistic reasoning are perfectly acceptable when viewed from a probabilistic perspective, that is, widespread intuitions are rendered more accurately by switching from logic to probability theory as a normative standard (Oaksford & Chater, 1994, 1996, 1998b, in press; Chater & Oaksford, 1999b).

Finally, Korb (2004) has also argued that a Bayesian approach could explain fallacies such as the appeal to authority and hence provides a framework for understanding ordinary arguments that is well worth developing.



The use of Bayesian probability to distinguish between warranted and unwarranted conclusions in the context of these fallacies is an attempt to develop a ‘reduction of fallacy theory’ in Lumer’s (2000) sense in that systematization and explanation of the fallacies is derived from a general normative theory. In contrast to past emphasis on deduction as the appropriate epistemological principle/standard, however, the Bayesian approach leads to a rather different evaluation of individual fallacies; the logical standard tends to lead to very ‘all or none’ evaluations whereas the probabilistic account allows and explains graded variation among instances of the same structure.

Because Bayesian probability provides a formal framework for distinguishing between warranted and unwarranted conclusions in the context of the fallacies, we think it has considerable potential for advancing epistemic approaches to argumentation (see e.g., Siegel & Biro, 1997; Goldman, 1997, 2003). Though we are optimistic in this regard, we do not wish to claim that the definitive, long-desired (see e.g., Hamblin, 1970) formal treatment of the fallacies has been provided in detail; for this, it will be necessary to demonstrate how Bayesian probability captures the bulk of the traditional list of fallacies (see Hahn & Oaksford, in press, for further discussion of the key issues here). Even less do we wish to claim that Bayesian probability has in any way been established as a sufficient theory of argument strength. For the current context we wish to claim only that some success in explaining the fallacies can be reported and that this success lends some credibility to the idea that a general, normative, Bayesian theory of argument strength might one day be forthcoming.

Clearly, this goal has not yet been achieved. What we wish to argue for in the remainder of this paper is why the pursuit of a specifically Bayesian theory of strength strikes us as worthwhile.

### **3. Why a *Bayesian* theory?**

In this section, we discuss some of the main assets of the Bayesian approach, which we think would make it particularly desirable as a normative theory.

Firstly, the Bayesian approach treats probabilities as expressions of subjective degree of belief not as an objective property of statements and their relationship to the world as is the case for frequentist interpretations of probability. This is important in the context of argumentation, because many of the things we argue about involve singular events—for example, whether or not Oswald killed JFK (see also Hahn & Oaksford, in press). Assigning single event probabilities only makes sense from a Bayesian subjective perspective; it is meaningless on a frequentist interpretation.<sup>2</sup> The fact that probabilities are taken to be expressions of subjective degree of belief also makes it particularly natural to interpret them as the degree of conviction associated with a claim. That conviction can be a matter of degree, not just a binary true or false, is fundamental for an adequate treatment of the fallacies. In



particular the treatment of circular arguments has been hampered by binary notions: Arguments involving self-dependent justification are frequent in science yet they are hopelessly rendered viciously circular and hence unacceptable by any account that conceives of statements put forward in an argument as only true or false (see Hahn, Oaksford & Corner, 2005 for detailed discussion; and more on this below). An all or none notion of dissent or assent to a claim has also led to an overexpansion of the notion of the burden of proof (see Hahn & Oaksford, *subm.*), suggesting, among other things, an explanation for why arguments from ignorance are poor that is vacuous in practice (see Hahn & Oaksford, *in press*; Hahn & Oaksford, *subm.* for detailed discussion).

Furthermore, the Bayesian formalism allows us to distinguish readily between the ultimate conviction with regards to a claim that an argument brings about—expressed as the posterior probability assigned to that claim in light of the evidence—and the *degree of change* that a reason effects (one way of measuring the latter is the likelihood ratio, see Hahn, Oaksford & Corner. 2005; Oaksford & Hahn, *in press*). This is important because the ultimate degree of conviction brought about in an argument is influenced, on the Bayesian account partly by the prior conviction associated with the claim. Consequently, we want to evaluate arguments not just with regards to how convinced they make us, but also with regards to how much they made us change our beliefs. This notion of degree of change (or ‘force’ of an argument) is important in the evaluation of argument strength, for example, because it allows one to explain why direct premise restatements (“God exists, because God exists”) make poor arguments even though they are deductively valid (see Hahn, Oaksford & Corner, 2005)—a tension that has puzzled philosophers for a long time. The ability to quantify change allows us to see clearly that such arguments bring about no change in convictions whatsoever. This makes them maximally ineffective as arguments and consequently maximally poor.<sup>3</sup>

That prior beliefs influence argument strength on the Bayesian account, of course, introduces a degree of relativity into the evaluation of arguments. We would argue that the degree of relativity afforded by the Bayesian approach is both essential and just right.

The important role a Bayesian analysis assigns to prior belief is an instance of a fundamental aspect of argumentation—the nature of the audience, which has been assumed to be a crucial variable for any rational reconstruction of argumentation (e.g., Perelman and Olbrechts-Tyteca, 1969; Goldman, 1997). Audience relativity has the consequence that a fallacy for one person may not be a fallacy for someone else because their prior beliefs differ. Ikenobe (2004) makes the same point using the argument that all cases of killing a living human being are bad and abortion is a case of killing a living human being, therefore, abortion is bad. This argument may provide adequate proof for someone who already believes that a fetus is a living human being. However, for someone who does not believe this proposition, this argument is weak or provides inadequate proof for the conclusion.

Crucially, however, the relativity introduced by Bayesian priors does not mean that “anything goes”. For one, the subjectivity introduced by priors does not mean that objectively true states cannot be reached. Individual Bayesian estimators can be more accurate than their frequentist counterparts, as judged by the criteria of the frequentist, even though they are biased estimators. For example, comparing Bayesian and frequentist estimators for a population mean, the Bayesian posterior mean will have the smaller error over the range of values that are realistic for that population mean (see, e.g., Bolstad, 2004). Furthermore, the well-known convergence properties of Bayesian updating mean that where enough suitable data are available, posteriors will eventually converge on the appropriate values regardless of priors. This latter result holds true, of course, only if the priors are not already ‘certain’, that is, 0 or 1. In this case, no amount of data can bring about further change. This allows one to capture the fact that some degree of “openness to change” or basic willingness to consider an argument is necessary for it to take an effect. At the same time it means that one does not lose the possibility of certainty associated with logical necessity and proof.

For many arguments in everyday life, of course, there will be neither proof nor huge amounts of data. Here, priors will matter and it should be seen as a virtue that they do. In the numerous argumentative contexts where there simply isn’t sufficient mutually agreed evidence to bring people’s beliefs in to alignment “with the facts” the Bayesian approach can reflect the diversity of opinion and legitimate disagreement that will remain.<sup>4</sup>

However, because Bayesian probability imposes constraints on the rational assignment of degrees of belief, the possibility of agreement and disagreement are constrained both within and across agents. The relative degree of conviction a set of different reasons brings about, for example, will be the same for two agents even where they differ in their priors with regards to the claim in question unless they also disagree about properties of the reasons themselves. In other words, even if we end up differentially convinced as a result of initial differences in priors, we can agree on the relative strength of arguments; where we do not do so, we must differ in other ways than just our priors for that disagreement to be rational.<sup>5</sup>

Moving on from the relativity (or not) afforded to argumentation by a Bayesian approach, a further important asset is the probabilistic notion of relevance (see also in the context of argument strength Korb; and, more generally, Pearl, 1988, 2000 for detailed discussion). ‘Conditional independence’ offers a dynamic notion of relevance, that changes as information is added or deleted to a database and the conditional independence axioms have been found to correspond to intuition about informational relevance in a variety of contexts (Pearl, 1988). Whether this is ultimately good enough, of course, remains to be seen. However, should more be required, some entirely new, as yet unknown account of relevance will likely have to be devised. It already seems comparatively clear that accounts of argumentative relevance that rely purely on logical consequence are unlikely to do justice to the concept of relevance required by a theory of informal argument. Similar arguments

have been made with respect to relevance logics (Anderson & Belnap, 1975). The concept of relevance goes beyond logical entailment, even *relevant* entailment. For example, it is relevant to whether an animal has palpitations that it has a heart but this is due to the causal structure of the world not the logical relations between propositions (Oaksford & Chater, 1991; Veltman, 1986). In short, existing theories of argumentation that transmit plausibilities via logical consequence relations seem unlikely to capture all the ways in which relevance relations are established in argument.

Our arguments in favor of a Bayesian approach so far have sought to identify characteristics that seem well attuned to the needs of argumentation. We conclude our survey of reasons why a specifically Bayesian theory of argument strength would be desirable with two more general considerations: the well-established normative standing of Bayesian probability and its connection with other bodies of research.

The most obvious consideration in developing a normative theory is the extent to which its normativity is indeed accepted or guaranteed. The normative standard of probability theory meets this requirement. There is an intuitive rational justification that underpins probability theory: Reasoning probabilistically is rational if one wants to avoid making bets one is guaranteed to lose. Furthermore, there is a well-defined formal calculus which guarantees that this rational principle is respected. The theory can deal with a huge range of hypotheses—whether these be discrete, multivalued, or continuous—and has been developed to deal with a wide variety of circumstances, such as uncertain or cascaded evidence. However, Bayesian conditioning, which is at the heart of the approach, follows directly from the three basic axioms of probability theory and the notion of conditional probability. That its assumptions are so minimal also lies at the heart of the finding that attempts to develop new and different formalisms frequently turn out to be ‘probabilities in disguise’ (on this issue see e.g., Cox, 1946; Horvitz, Heckerman & Langlotz, 1986; Heckerman, 1986; Snow, 1998; see also, Pearl, 1988 and Howson & Urbach, 1993 for further references). Consequently, we follow Pearl’s (1988, p. 20) view on whether it is necessary to supplant probability theory, “...we find it more comfortable to compromise an ideal theory [i.e., probability theory] that is well understood than to search for a new surrogate theory, with only gut feeling for guidance.”

The final benefit of the Bayesian approach then, is that it connects research on everyday argument with a number of different bodies of research. At the level of normative theory, our Bayesian approach trades on similar approaches to scientific inference (e.g., Howson & Urbach, 1993; Earman, 1992), which, following other authors in the area (Kuhn, 1993), we have suggested can be extended naturally to informal argument. This is advantageous not only because of the potential theoretical unification, but also because a lot of hard work has already been done. The Bayesian approach to scientific inference has received much scrutiny and criticism (e.g., Miller, 1994; Sober, 2002); not all of this is equally relevant to everyday argument

(specifically, the subjectivity inherent in the Bayesian approach might, to some, be more palatable in non-scientific contexts), but a wealth of important issues have already been well worked through and can be received into studies of argumentation. One example of this is the issue of priors and ignorance that we will return to below.

Within psychological research on human behavior, finally, the Bayesian approach links up with the ever-increasing body of evidence that suggests that much of human cognition is, in one way or the other, probabilistic. Theories and data range from Bayesian accounts of vision (Knill & Whitman, 1996), through language acquisition and processing (e.g., Bates & MacWhinney, 1989; MacDonald, 1994), to many aspects of higher level cognition (e.g., Oaksford & Chater, 1998b). An emerging picture, here, suggests that humans are probabilistic beings.<sup>6</sup> A probabilistic account of argument evaluation receives support from this ‘character’ argument.

#### **4. Limitations of a Bayesian Approach**

The intention in this paper is to highlight the merits of a Bayesian approach, not a comparative evaluation with possible competitors, not least because that set itself is not closed. However, it would be wrong to complete a discussion of a Bayesian approach to argumentation without drawing attention to the fact that Bayesian principles have been subject to criticism. Because the interest in Bayesian principles in areas such as statistics and also the philosophy of science has been intense, these criticisms have been well-discussed and many issues have more or less well-developed replies. For an excellent overview the interested reader is referred to Howson and Urbach’s volume on scientific reasoning (1993). We restrict ourselves here to two issues that have found their way into the literature on argumentation.

First is the idea that what we know, or more importantly do not know, makes the Bayesian formalism unnatural. Walton (2004, pg. 277) and Ennis (2004), for example, seem to echo an oft heard claim that the assignment of numerical values to premises is frequently impossible or unhelpful, in that it requires an exactness that is not available in most cases.

This slightly misses the point of subjective probabilities: they are *expressions*—through the use of non-extreme values—of the indefiniteness of one’s knowledge and are introduced precisely because of uncertainty. The argument, by contrast, makes it seem as if one not only has to know that one is uncertain, but also that one has to know to a precise degree how uncertain (see also Howson & Urbach, 1993 pg. 87). While it is true that some number has to be specified, the resolution with which that number is specified can vary according to context and sensitivity analysis can be used to determine how much precision the problem at hand requires (on sensitivity analysis see e.g., Gill, 2002). Exactly the same applies to the use of real numbers in measuring physical, everyday quantities. In many contexts, it will simply be irrelevant to the required outcome not only whether a measurement was 5.687 or 5.689cm but also whether it was 5.5 or 5.7cm, or even 5 or 6cm. In

practice, one will *always* be restricting oneself to certain points of the scale, and what these are will be determined by context.

Likewise, even though a specific point value probability has to be specified, there is nothing to stop one, in a given context, restricting one's use of the scale to 5 different points, for example 0.1, 0.25, 0.5, 0.75, 0.9 which one might take to correspond to the verbal descriptors 'very unlikely', 'unlikely', '50/50 chance', 'likely', and 'almost certain'—or any other such set of numbers by which to express broad increments. Alternatively, one can define and use *intervals* over the scale. As long as these intervals are small, one gets a generalization of the classic calculus which is simply defined over interval values rather than point values (see e.g., Walley, 1991; for a broader overview and discussion see also Parsons, 2001). It is also interesting to note here that peoples' attitudes toward imprecision in the specification of probabilities seems to contain an asymmetry. In experimental contexts it has been found that while people prefer to express probabilities with verbal descriptors and the crude distinctions they afford, they actually prefer to *receive* numerical probabilities (e.g., Wallsten *et al.*, 1993); so it does not actually seem to be the case that numerical expressions of probability are inherently 'unnatural'.

The issue of indefiniteness in one's degree of uncertainty leads on naturally to the case of maximal indefiniteness in the form of complete ignorance. Because the Bayesian treatment of complete ignorance is one of the most widely cited 'shortcomings' of Bayesian probability, it seems useful to also discuss it here. The Bayesian way to represent ignorance about a range of possibilities is to assign them all equal probability according to the 'Principle of Indifference'. This principle states that when we have a number of possibilities, with no relevant difference between them, they all have the same probability. As a way of representing ignorance, indifference makes intuitive sense; however, it readily leads to seeming 'paradox'. As an example, one might take an urn filled with white and colored balls in unknown proportions, but the colored balls consist of red balls and blue balls in equal number. According to the Principle of Indifference, our present data—before we have drawn our first ball—are neutral between its being colored and its being white. Hence, we should expect it to be white with a probability of .5. However, if the ball is colored, then it is either red or blue, and the data are also neutral between the ball's being either white or red or blue. Hence, according to the Principle of Indifference, the probability of the ball's being white is one third. Another example, this time concerning a garden plot, is presented by Sober (2002). Once again there is no unique way to translate ignorance into an assignment of priors: one gets one answer if one applies a uniform prior to lengths of sides of plot and another if one applies it to the area of the same plot. Examples of this kind can be generated ad nauseam (see Howson & Urbach, 1993 for further examples and references). What then, are their implications?

The paradoxes arise because the uniform assignment of probabilities is language- or description relative. It is the primitives of the description that are assigned equal

probability according to the Principle of Indifference. Terms *derived* from these primitives will not themselves necessarily have equal probability. Moreover, they will not do so for good mathematical or logical reason. If all six numbers of a dice are equi-probable, then the outcome ‘greater than 1’ will have a 5 in 6 chance of occurring. The seeming paradox occurs, because alternative sets of primitives give rise to different probabilities for composite beliefs—but given our ignorance, the choice between these sets seems arbitrary.

Some have seen these paradoxes as so compelling as to force the development of alternative theories. Dempster-Shafer theory, for example, is based on the idea that the natural way to represent total ignorance is to treat equally all possible alternatives, whether they are primitive or composite. This would allow one to have the same degree of belief both in the claim that all six numbers are equally likely *and* that ‘greater than 1’ is as likely as not to occur (see Howson & Urbach, 1993).

The Bayesian response to this is that total ignorance is not possible. Specifically it is not *rationally* possible to be uniformly unopinionated about everything. The probability calculus provides a normative theory for the rational assignment of belief. From certain beliefs other beliefs will necessarily follow, for example, by logical consequence. A rational agent should be committed to these beliefs whether or not they are held in actual fact or even entertained. There are considerable constraints on what is a rational set of beliefs given no evidence. If I actively—in ignorantiam—believe something about a hypothesis  $H$ , then I am forced to believe something else about  $\neg H$ ; likewise, if I actively—in ignorantiam—believe something about the perimeter of a property, I have to consistently believe something about its area. As Howson and Urbach argue, it should be seen as a virtue not a vice of the theory that it brings out the impossibility of total ignorance so clearly.

The argument based on paradoxes seems compelling because it is so easy to be unaware of these constraint—whether this is due to a failure of rationality, or a reflection of the fact that one actually held no beliefs on the topic whatsoever. The fact that different problem statements can give rise to different answers and that this discrepancy might seem arbitrary if there is not much to choose between them should likewise not be seen as a fault of the formalism. That different answers can and will ensue depending on how a formalism is mapped onto a real world problem is not a unique property of Bayesian inference, but occurs wherever states of affairs are mapped onto formal models. This problem could be avoided only if there was one single, unique way to represent the world. In lieu of that, the problem will remain, and it is typically *not* the job of the formalism itself to determine which of many possible or even plausible mappings is preferable. What is comforting about Bayesian inference and the so-called ‘paradoxes’ involving ignorance is that it need not always matter: once balls are drawn from the urn my estimate will gradually become more and more accurate, whether my prior was  $1/2$  or  $1/3$ .

The second perceived limitation of the Bayesian approach that can be found in the argumentation literature concerns the way argument strength is calculated



from evidence. An alternative approach to argumentation views arguments as *presumptively* acceptable (e.g., Walton, 1996). That is, they are *defaults* that can be readily overturned by new information. Walton makes the direct link between this style of reasoning and the development of *non-monotonic* logics (e.g., Reiter, 1980, 1985) in AI. In AI knowledge representation, the failure to develop tractable non-monotonic logics that adequately capture human inferential intuitions, has provided part of the impetus for the development of Bayesian probabilistic approaches to uncertain reasoning. Nonetheless, some alternative approaches have been developed that explicitly include some measure of the strength of an argument (Gabbay, 1996; Fox & Parsons, 1998; Pollock, 2001; Prakken & Vreeswijk, 2002). Moreover, it is this inclusion that provides these systems with their nice default properties.

These systems explicitly eschew the idea that argument strength can be adequately dealt with using the probability calculus. The root of this contention is *Theophrastus' Rule*: the strength that a chain of deductively linked arguments confers on the conclusion cannot be weaker than the weakest link in the chain (Walton, 2004). This is a condition that cannot be guaranteed by the probability calculus. Examples that seem to conform to Theophrastus' rule but not to the probability calculus, have persuaded Walton (2004) and, for example, Pollock (2001), that a third form of reasoning should be countenanced in addition to deductive and inductive/statistical reasoning, i.e., *plausibilist* reasoning.

We address first one of the examples for which probability has been viewed as inadequate. We then address directly Theophrastus' rule.

Walton's (2004) example concerns negation within the probability calculus according to which the probability of the negation of a hypothesis  $H$ ,  $P(\neg H)$ , is constrained to be 1 minus the probability of that hypothesis,  $P(H)$ . Plausibilist reasoning based on Theophrastus' rule is necessary because this probabilistic approach to negation is sometimes inappropriate. Specifically, Walton argues that in the case of legal argumentation involving evidence, both a proposition and its opposite can be highly plausible. He writes,

For example, suppose a small and weak man accuses a large and strong man of assault. The small man argues that it is implausible that he, the weaker man, would attack a visibly stronger man who could obviously defeat him. In court, the visibly larger and stronger man asks whether it is plausible that he would attack such a small man in front of witnesses when he knew full well that he could be accused of assault. Here we have an argument with probative weight on one side, but also an argument with probative weight on the opposed side. The proposition that the large man committed assault is plausible in light of the facts, but its negation is also plausible in light of the same facts. The argumentation in this typical kind of case in law violates the negation axiom of the probability calculus. (pg. 278)

From this it follows, according to Walton, that a different calculus is necessary, namely one whereby a proposition can be plausible in relations to a body of evidence



in a given case, whereas the negation of that same proposition can also be plausible in relation to another body of evidence, and that a contextual and pragmatic notion of probative evidence is required.

We do not think the example warrants a new form of reasoning. What is going on here is simply a case of conflicting evidence. One first receives evidence that would increase one's posterior degree of belief in the claim, and then receives evidence that would decrease it. In this sense, one has had evidence both to make the claim more plausible, and evidence to reduce its plausibility. However, argument evaluation *must* mean that the two sources of evidence are at some point integrated into a final, overall assessment. At this stage, they can either cancel each other out, or one can outweigh the other with the consequence that one's final conviction is shifted in the direction of the 'weightier' evidence, though less so than if the counterevidence had not been received. One does not *end up* both more convinced that the small man hit the large one and that he did not.<sup>7</sup> Unless one views both bodies of evidence as entirely equal, in which case nothing changes, weighing the evidence will increase one's belief in the one at the expense of the other. This is exactly what the probability calculus allows one to achieve. The first set of evidence increases one's posterior degree of belief in the claim, the second decreases it and Bayesian updating provides the mechanism whereby these beliefs are integrated into an overall single judgment whether the two sets of evidence are received one after the other or together.

It is also important here that not all legal claims involve a proposition and its negation. They might equally involve two propositions that are mutually exclusive but not complements. In fact, Walton's example seems at times to oscillate between negation and mutual exclusivity in that the small man's claim that it would be implausible that he would attack a larger man seems more relevant to a debate about who threw the first punch as opposed to a debate about whether or not the large man hit him. In arguing about the first punch, the answer could be the small man, the large man, both men simultaneously or that nobody hit anybody at all. In this context, the small man's claim that the large man hit him first could become more plausible according to the evidence as could the large man's claim that the small man hit him first, because only the *sum total* of the probabilities associated with the four logical possibilities must equal to one. In other words, one could become more convinced that one of them hit the other, as opposed to no fight having taken place at all, but be none the wiser as to which one hit the other first. Again, Bayesian probability will give this result without the need for a new form of 'plausible reasoning'.

To conclude our discussion, we provide an example to illustrate why we think Theophrastus' rule is not a good idea. The example is drawn from our treatment of the fallacies and involves circular arguments. Most circular arguments found in practice do not involve a direct restatement of the conclusion among the premises, rather the conclusion forms an implicit assumption, a presupposition, that underlies the interpretation of the premise material. Hence, the inference involves self-dependent

justification (also referred to as ‘epistemic circularity’, e.g., Goldman, 2003; see Hahn, Oaksford & Corner, 2005 for discussion of the previous literature).

This is the case for the classic textbook example,

- (7) God exists, because the Bible says so, and the Bible is the word of God.

However, this is also the case for a huge number of scientific inferences, for example

- (8) Electrons exist, because we can see 3 cm tracks in a cloud chamber, and 3 cm tracks in cloud chambers are signatures of electrons.

The scientific example which arises wherever scientists are dealing with entities that cannot be directly observed seems a perfectly acceptable example of the classic *inference to the best explanation* (Harman, 1965; see also, Josephson & Josephson, 1994). This suggests that self-dependent justification is not inherently wrong (see also Shogenji, 2000 and Brown, 1993, 1994 for scientific examples). Our Bayesian account explains why (and for whom) the Bible example, which has exactly the same structure, seems weak and the scientific example acceptable.

Self-dependent justification of this kind is captured through hierarchical Bayesian inference. Three levels are involved here—the directly observed evidence, the interpretation of the evidence, and the hypothesis. The observed evidence is the Bible’s claim that God exists and the 3cm tracks seen in the cloud chamber. These are relevant to the hypothesis because they are interpreted as the word of God and the signature effects of electrons, respectively. This interpretation is itself uncertain and dependent on the fact that the hypothesis is true. This extra level, however, does not in anyway preclude Bayesian conditioning on the given observation (see e.g., Pearl, 1988 for examples of hierarchical Bayesian inference), and making that observation will increase our posterior degree of belief in the hypothesis. In the case of the Bible, this increase will typically be slight, because there are numerous other plausible interpretations of the Bible and its content other than that it is the direct word of God, and priors, as always, also affect how convincing the argument will be. The scientific example will seem stronger simply because (and as long as) our estimates of the associated probabilities are different. A Bayesian analysis then explains the difference between weak, textbook examples of self-dependent justification and widespread scientific practice and renders scientific use of inference to the best explanation acceptable.

By contrast, classical logic fails here (Hahn, Oaksford & Corner, 2005 for fuller discussion). Because the conclusion must already be assumed as a premise, and degrees of belief (i.e., ‘truth’) are all-or-none, no self-dependent argument can bring about any change in conviction. Consequently, all such arguments are necessarily rendered maximally poor.

Moreover, Theophrastus’ rule fails here as well. The interpretation of the observation statement required for the inference (i.e., that 3cm tracks are signature

effects of electrons; for example, random smudges on the screen) depends on the hypothesis in question being true (i.e., trivially, if there is no such thing as an electron the 3cm tracks can be whatever they like except an electron's signature effect). This dependency means that the probability of the interpretation being true *can never be greater* than the probability of the hypothesis itself. This follows directly from the fundamental axioms of probability theory, because the interpretation statement, if true, would imply the truth of the hypothesis, and the probability of any logical consequence of a statement must be at least as great as the probability of that statement itself.

In other words, the degree of belief associated with the interpretation statement will be a weaker (or as weak a) link in the premise material as the presupposed conclusion itself. Hence our degree of belief in that conclusion should *never* rise according to Theophrastus' rule. Much of our scientific reasoning and argumentation would be labelled 'fallacious' as a result, and my degree of belief in electrons would remain the same regardless of what I observed in the cloud chamber experiment or, in fact, whether or not I bothered to conduct the experiment at all.

This example, in our view, also serves to underscore the wider merit of sticking with a well-established normative calculus in that trouble free inference procedures are not that easy to derive, let alone to establish as normatively justified.

## 5. Conclusion

We have sought to argue here for the desirability not only of normative theories of argument strength, but also for the desirability of a specifically Bayesian account. Bayesian probability brings with it a whole host of inherent characteristics and theoretical connections that make it an attractive candidate framework. Also, its success in explaining a range of fallacies suggests, to us at least, that there is some chance its promise might be fulfilled. It is hoped that the case for Bayesian probability made here will motivate future research aimed at this goal.

## Acknowledgments

The authors would like to thank Adam Corner and Adam Harris for helpful discussions, as well as Christoph Lumer and three anonymous reviewers for their useful comments on this manuscript.

## Notes

<sup>1</sup> A cash register, as a simple computing device, calculates a total sum of cost (computational level description) as part of the social exchange network that constitutes a 'sale'- several possible algorithms for addition are available as procedures for the device at the representational level, and each of these, finally, could be realized in a physical device in countless ways.

<sup>2</sup> Propensity theory (e.g., Popper, 1959) is another suggested interpretation of the probability calculus which conceptually allows single-event probabilities, while seeking to maintain a frequentists basis. On its problems see e.g. Howson & Urbach, 1993, ch. 13 for detailed discussion.

<sup>3</sup> Note that this is by no means a property of *all* logically valid arguments (Hahn et al, 2005). It does not apply, for example, to the inference  $p$  therefore ' $p$  or  $q$ '.

<sup>4</sup> Ennis (2004) claims that assigning subjective probabilities to statements altogether 'wipes out disagreement' among people and hence is inappropriate for argumentation, because we can agree that our subjective views of a claim differ. However, we do not see how it follows from this that we cannot be motivated to *change* others' degree of belief to the extent that it does not correspond to our own, so argument is still both meaningful and possible.

<sup>5</sup> Specifically the likelihood ratio has to be different.

<sup>6</sup> This, of course, does not imply that peoples assessments of probabilities are always entirely accurate. It does mean, however, that sensitivity to the probabilistic nature of the environment emerges as a central aspect of computational level description (see section 1). At the same time, however, subsequent research has greatly modified some of the early claims regarding people's 'failings' with regards to intuitive statistics (e.g., Birnbaum 2004, cf. Tversky & Kahneman, 1974).

<sup>7</sup> If one did, one would be prone to (synchronic) Dutch books, that is bets one is guaranteed to lose. Specifically, in the case where it is clear that either the A hit B or B hit A, becoming both more convinced in A's hitting and in B's hitting would mean that the degrees of belief assigned to each of these possibilities could exceed 1. Assume one thought, for example, that there was a .6 chance that A hit B, and a .5 chance that it was B that hit A. Assume further one is willing to bet in line with one's degrees of belief, such that one will accept up to or equal to one's 'degree of belief' x \$10 for a unit wager that pays \$10 if the claim in question turns out to be true. One would then be happy to pay a bookie \$6 on A being the hitter, but also to pay \$5 on B being the hitter (i.e., 'not A'). Regardless of who had actually done the hitting, one would then lose \$1, having paid \$11 on a combination of wagers guaranteed to pay exactly \$10. By contrast, as long as degrees of belief in a statement and its complement sum to one, bets on an event and its complement respecting those degrees of belief will break even.

## References

- Alexy, R. (1989). *A Theory of Legal Argumentation*. Oxford: Clarendon Press.
- Allen M. & Kellermann, K. (1988). Using the subjective probability model to evaluate academic debate arguments. *Argumentation and Advocacy*, 25, 93-107.
- Allen, M., Burrell, N. & Egan, T. (2000) Effects with multiple causes: evaluating arguments using the subjective probability model. *Argumentation & Advocacy*, 37, 109-116.
- Anderson, J. (1991). Cognitive Architectures in a rational analysis. In K. VanLehn (ed.), *Architectures for Intelligence*, pp. 1-24. Hillsdale, NJ: Erlbaum.
- Anderson, A.R., & Belnap, N.D. (1975). *Entailment: The logic of relevance and necessity Vol. 1*. Princeton: Princeton University Press.
- Bates, E. & MacWhinney, B. (1989). Functionalism and the competition model. In E. Bates, & B. MacWhinney (Eds.), *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.
- Birnbaum, M. H. (2004). Base rates in bayesian inference. In R. F. Pohl (Ed.), *Cognitive Illusions* (pp. 43-60). New York, NY: Psychology Press.
- Boger, G. (2005). Subordinating truth—is *acceptability* acceptable? *Argumentation*, 19, 187-238.
- Bolstad, W.M. (2004). *Introduction to Bayesian Statistics*. Hoboken, NJ: Wiley.

- Bowell, T. & Kemp, G. (2002). *Critical thinking: A concise guide*. London: Routledge.
- Brown, H. (1993). A theory-laden observation *can* test a theory. *British Journal for the Philosophy of Science*, 44, 555-559.
- Brown, H. (1994). Circular justifications. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1, 406-414.
- Chater, N., & Oaksford, M. (1999a). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, 3, 57-65.
- Chater, N., & Oaksford, M. (1999b). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, 38, 191-258.
- Chater, N., & Oaksford, M. (2000). The rational analysis of mind and behaviour. *Synthese*, 122, 93-131.
- Corner, A., Hahn, U. & Oaksford, M. (2006) The slippery slope argument: probability, utility and category reappraisal. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*.
- Cox, R. (1946). Probability frequency and reasonable expectation. *American Journal of Physics*, 14, 1-13.
- Davidson, D. (1984). On the very idea of a conceptual scheme. In *Inquiries into Truth and Interpretation* (pp. 183-198). Oxford: Oxford University Press.
- Earman, J. (1992). *Bayes or bust?* Cambridge, MA: MIT Press.
- Eemeren, F.H. van, & Grootendorst, R. (1992). *Argumentation, communication, and fallacies*. Hillsdale, NJ: Lawrence Erlbaum.
- Eemeren, F.H. van, & Grootendorst, R. (2004). *A systematic theory of argumentation. The pragma-dialectical approach*. Cambridge: Cambridge University Press.
- Ennis, R.H. (2004). Applying soundness standards to qualified reasoning. *Informal Logic*, 24, 23-39.
- Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skill. *Discourse Processes*, 32, 135-153.
- Feteris, E.T. (1997). A survey of 25 years of research on legal argumentation. *Argumentation*, 11, 355-376.
- Fox, J., & Parsons, S. (1998). Arguing about beliefs and actions. In A. Hunter and S. Parsons (Eds.), *Applications of uncertainty formalisms (Lecture notes in artificial intelligence 1455)* (pp. 266-302). Berlin: Springer Verlag.
- Freedman, J. & Fraser, S. (1966). Compliance without pressure: the foot in the door technique. *Journal of Personality and Social Psychology*, 4, 195-202.
- Gabbay, D. (1996). *Labelled deduction systems*. Oxford: Oxford University Press.
- Gill, J. (2002) *Bayesian methods—a social and behavioural sciences approach*. Boca Raton, FL: Chapman & Hall/CRC.
- Goldman, A.I. (1994) Argumentation and social epistemology. *The Journal of Philosophy*, 91, 27-49.
- Goldman, A.I. (1997). Argumentation and interpersonal justification. *Argumentation*, 11, 155-164.
- Goldman, A.I. (2003). An epistemological approach to argumentation. *Informal Logic*, 23, 51-63.

- Hahn, U. & Oaksford, M. (in press). A Bayesian approach to informal reasoning fallacies. *Synthese*.
- Hahn, U. & Oaksford, M. (subm.). *The burden of proof and its role in argumentation*.
- Hahn, U., Oaksford, M., & Bayindir, H. (2005). How convinced should we be by negative evidence? In B. Bara, L. Barsalou, and M. Bucciarelli (Eds.), *Proceedings of the 27<sup>th</sup> Annual Conference of the Cognitive Science Society*, (pp. 887-892), Mahwah, N.J.: Lawrence Erlbaum Associates.
- Hahn, U., Oaksford, M., & Corner, A. (2005). 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.
- Hamblin, C.L. (1970). *Fallacies*. London: Methuen.
- Hample, D. (1977) Testing a model of value, argument, and evidence. *Communication Monographs*, 44, 106-120.
- Hample, D. (1978) Predicting immediate belief change and adherence to argument claims. *Communication Monographs*, 45, 219-228.
- Hample, D. (1979) Predicting belief and belief change using a cognitive theory of argument and evidence. *Communication Monographs*, 46, 142-151.
- Harman, G. (1965). The inference to the best explanation. *Philosophical Review*, 64, 88-95.
- Heckerman, D. (1986). A rational measure of confirmation. In L.N. Kanal and J.F. Lemmer (eds.) *Uncertainty in Artificial Intelligence 2*. Amsterdam: North Holland.
- Heyse, T. (1997). Why logic doesn't matter in the (philosophical) study of argumentation. *Argumentation*, 11, 211-224.
- Holyoak, K. J. & Morrison, R.G. (2005). *The Cambridge Handbook of Thinking and Reasoning*. Cambridge: Cambridge University Press.
- Horvitz, E.J., Heckerman, D., & Langlotz, C.P. (1986) A framework for comparing alternative formalisms for plausible reasoning. In, *Proceedings of the 5<sup>th</sup> National Conference on AI (AAAI-1986)*, pg 210-214.
- Howson, C., & Urbach, P. (1993) *Scientific reasoning: The Bayesian approach, 2<sup>nd</sup> edition*. La Salle, IL: Open Court.
- Ikuenobe, P. (2004). On the theoretical unification and nature of the fallacies. *Argumentation*, 18, 189-211.
- Johnson, B.T., Maio, G.R. & Smith-McLallen (in press) Communication and Attitude Change: Causes, Processes and Effects. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes and attitude change: Basic principles*. Mahwah, NJ: Erlbaum.
- Johnson, R.H. (2000). *Manifest rationality: a pragmatic theory of argument*. Mahwah, NJ: Erlbaum.
- Josephson, J.R., & Josephson, S. G. (1994). *Abductive inference: Computation, philosophy, technology*. Cambridge: Cambridge University Press.
- Kahneman, D., Slovic, P. and Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

- Knill, D.C. & Whitman, R. (eds.) (1996). *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.
- Kohlas, J. (2003). Probabilistic argumentation systems: A new way to combine logic with probability. *Journal of Applied Logic*, 1, 225-253.
- Korb, K. (2004). Bayesian Informal Logic and Fallacy. *Informal Logic*, 24, 41-70.
- Kuhn, D. (1993). Connecting scientific and informal reasoning. *Merrill-Palmer Quarterly*, 39, 74-103.
- Lode, E. (1999). Slippery slope arguments and legal reasoning. *California Law Review*, 87, 1469-1544.
- Lumer, C. (1990) *Praktische Argumentationstheorie. Theoretische Grundlagen, praktische Begründung wichtiger Argumentationsarten*. Braunschweig: Vieweg.
- Lumer, C. (1997). Practical arguments for theoretical theses. *Argumentation*, 11, 329-340.
- Lumer, C. (2000). Reductionism in fallacy theory. *Argumentation*, 14, 405-423.
- MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, 9(2), 157-201.
- Maio, G. R., & Haddock, G. (in press). Attitude change. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social Psychology: Handbook of Basic Principles* (Vol. 2). New York, NY: Guilford.
- Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman and Company.
- McGuire, M. (1960). Cognitive consistency and attitude change. *Journal of Abnormal and Social Psychology*, 60, 345-353.
- McPeck, J. (1981). *Critical thinking and education*. Oxford: Martin Robertson.
- Miller, D. (1994). *Critical rationalism: A restatement and a defence*. La Salle, IL: Open Court.
- Neuman, Y. (2003). Go ahead, prove that God does not exist! *Learning and Instruction*, 13, 367-380.
- Neuman, Y., & Weitzman, E. (2003). The role of text representation in students' ability to identify fallacious arguments. *Quarterly Journal of Experimental Psychology*, 56A, 849-864.
- Neuman, Y., Weinstock, M.P., & Glasner, A. (in press). The effect of contextual factors on the judgment of informal reasoning fallacies. *Quarterly Journal of Experimental Psychology*
- Neumann, U. (1986). *Juristische Argumentationslehre*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Oaksford, M. & Chater, N. (1991). Against logicist cognitive science. *Mind & Language*, 6, 1-38.
- Oaksford, M. & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608-631.
- Oaksford, M. & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review*, 103, 381-391.
- Oaksford, M. & Chater, N. (Eds.) (1998a). *Rational models of cognition*. Oxford: Oxford University Press.
- Oaksford, M. & Chater, N. (1998b). *Rationality in an uncertain world: Essays on the cognitive science of human reasoning*. Hove, Sussex: Psychology Press.



- Oaksford, M. & Chater, N. (in press). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.
- Oaksford, M., & Hahn, U. (2004). A Bayesian approach to the argument from ignorance. *Canadian Journal of Experimental Psychology*, 58, 75-85.
- Oaksford, M. & Hahn U. (in press) Induction, deduction and argument strength in human reasoning and argumentation. In A., Feeney & E., Heit (Eds.) *The psychology of induction*. Cambridge: Cambridge University Press.
- Osherson, D.N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, 97, 185-200.
- Parsons, S. (2001). *Qualitative methods for reasoning under uncertainty*. Cambridge, MA: MIT Press.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan Kaufman.
- Pearl, J. (2000). *Causality*. Cambridge: Cambridge University Press.
- Perelman, C., & Olbrechts-Tyteca, L. (1969). *The new rhetoric: A treatise on argumentation*. Notre Dame, IN: University of Notre Dame Press.
- Pollock, J.L. (2001). Defeasible reasoning with variable degrees of justification. *Artificial Intelligence*, 133, 233-282.
- Popper, K.R. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science*, 10, 25-42.
- Prakken, H., & Vreeswijk, G.A.W. (2002). Logics for defeasible argumentation. In D.M. Gabbay and F. Guenther (Eds.), *Handbook of Philosophical Logic, 2nd edition, Vol 4* (pp. 219-318). Dordrecht/Boston/London: Kluwer Academic Publishers.
- Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, 13, 81-132.
- Reiter, R. (1985). On reasoning by default. In R. Brachman, & H. Levesque (Eds.) *Readings in knowledge representation* (pp. 401-410). Los Altos, CA: Morgan Kaufman.
- Rips, L.J. (1998). Reasoning and conversation. *Psychological Review*, 105, 411-441.
- Rips, L.J. (2001). Two kinds of reasoning. *Psychological Science*, 12, 129-134.
- Rips, L.J. (2002). Circular reasoning. *Cognitive Science*, 26, 767-795.
- Savage, L.J. (1954). *The foundations of statistics*. New York: John Wiley.
- Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and education*. New York: Routledge.
- Siegel, H. & Biro, J. (1997). Epistemic normativity, argumentation, and fallacies. *Argumentation*, 11, 277-292.
- Shogenji, T. (2000). Self-dependent justification without circularity. *British Journal for the Philosophy of Science*, 51, 287-98.
- Slob, W.H. (2002) How to distinguish good and bad arguments: dialogico-rhetorical normativity. *Argumentation*, 16, 179-196.
- Snow, P. (1998). On the correctness and reasonableness of Cox's Theorem for finite domains. *Computational Intelligence*, 14, 452-459.
- Sober, E. (2002). Bayesianism: Its scope and limits. In R. Swinburne (Ed.), *Bayes theorem*. Oxford: Oxford University Press.

- Tversky, A. & Kahneman, D. (1974) Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Veltman, F. (1985). *Logics for conditionals*. Ph.D. Thesis, Faculteit der Wiskunde en Natuurwetenschappen, University of Amsterdam.
- Walley, P. (1991) Statistical reasoning with imprecise probabilities. London: Chapman and Hall.
- Wallsten, T.S., Budescu, D.V., Zwick, R. & Kemp, S.M. (1993). Preferences and reasons for communicating probabilistic information in numerical or verbal terms. *Bulletin of the Psychonomic Society*, 31, 135-138.
- Walton, D.N. (1992). Nonfallacious arguments from ignorance. *American Philosophical Quarterly*, 29, 381-387.
- Walton, D.N. (1995). *A pragmatic theory of fallacy*. Tuscaloosa/London: The University of Alabama Press.
- Walton, D.N. (1996). *Arguments from ignorance*. Philadelphia: Pennsylvania State University Press.
- Walton, D.N. (2004). *Relevance in Argumentation*. Mahwah, NJ: Erlbaum.
- Weinstock, M., Neuman, Y. & Tabak, I. (2004). Missing the point or missing the norm? Epistemological norms as predictors of students' ability to identify fallacious arguments. *Contemporary Educational Psychology*, 29, 77-94.
- Williams, B. (1985). Which slopes are slippery? In M. Lockwood (Ed.), *Moral dilemmas in modern medicine* (pp. 126-137). Oxford: Oxford University Press.
- Woods, J., Irvine, A., & Walton, D.N. (2004). *Argument: Critical thinking, logic and the fallacies, Revised Edition*. Toronto: Prentice Hall Canada.
- Wyer, R.J. (1970) Quantitative prediction of belief and opinion change: a further test of a subjective probability model. *Journal of Personality and Social Psychology*, 16, 559-570.
- Wyer, R. (1974). Some implications of the "Socratic effect" for alternative models of cognitive consistency. *Journal of Personality*, 42, 399-419.
- Wyer, R. & Goldberg, L. (1970). A probabilistic analysis of the relationships among beliefs and attitudes. *Psychological Research*, 77, 100-120.

*Ulrike Hahn*  
School of Psychology  
Cardiff University  
Tower Building, Park Place  
Cardiff CF10 3AT, U.K.

*hahnu@cardiff.ac.uk*

*Mike Oaksford*  
School of Psychology  
Birkbeck College London  
Malet Street  
London, WC1E 7HX  
United Kingdom

*m.oaksford@bbk.ac.uk*