Synthese (2013) 190:3579–3610
DOI 10.1007/s11229-012-0211-y

Normative theories of argumentation: are some norms better than others?

Adam Corner · Ulrike Hahn

Received: 21 May 2012 / Accepted: 18 October 2012 / Published online: 3 November 2012
© Springer Science+Business Media Dordrecht 2012

Abstract Norms—that is, specifications of what we ought to do—play a critical role in the study of informal argumentation, as they do in studies of judgment, decision-making and reasoning more generally. Specifically, they guide a recurring theme: are people rational? Though rules and standards have been central to the study of reasoning, and behavior more generally, there has been little discussion within psychology about why (or indeed if) they should be considered normative despite the considerable philosophical literature that bears on this topic. In the current paper, we ask what makes something a norm, with consideration both of norms in general and a specific example: norms for informal argumentation. We conclude that it is both possible and desirable to invoke norms for rational argument, and that a Bayesian approach provides solid normative principles with which to do so.

Keywords Norms · Argumentation · Reasoning · Bayesian probability

1 Introduction: norms and human cognition

In the increasingly popular study of informal argumentation, norms play a critical role—and two broad sets of normative theories can be distinguished. *Procedural* theories of argumentation seek to specify rules of engagement for argumentative practice, that is, the procedural rights and obligations of those taking part in rational debate (e.g., van Eemeren and Grootendorst 2004). By contrast, *epistemic* theories focus on the actual truth or falsity of argumentative claims (e.g., Biro and Siegel 2006; Hahn and Oaksford 2007) and seek to identify appropriate standards such as logic or probability

A. Corner (✉)
School of Psychology, Cardiff University, Tower Building, Park Place Cardiff, CF10 3AT, UK
e-mail: corneraj@cardiff.ac.uk

U. Hahn
Department of Psychological Sciences, Birkbeck, University of London, London WC1E 7HX, UK

theory with which to assess people's use and acceptance of different types of argument. Normative considerations play a critical role in theories of fallacy—the supposedly faulty forms of inference that fill guides to critical thinking (see e.g., Woods et al. 2004)—as well as empirical tests of informal argument evaluation (Bailenson and Rips 1996; Brem and Rips 2000; Corner and Hahn 2009; Corner et al. 2011; Hahn and Oaksford 2006a,b, 2007; Hahn et al. 2005; Hoeken 2001a,b; Neuman et al. 2006; Oaksford and Hahn 2004; Rips 1998, 2001; Siegel and Biro 1997; van Eemeren and Grootendorst 2004).

Norms are also invoked in the many programmes designed to assess the development of argument skills in children (Driver et al. 2000; Kuhn and Udell 2003; Simon et al. 2002), or to improve the arguments used in science communication (Corner and Hahn 2009; Korpan et al. 1997; Norris et al. 2003). Norms have provided an important framework for identifying discrepancies between the goals of advocates and the actual persuasive effect of their messages (O'Keefe 2003, 2007). And despite pursuing an ostensibly descriptive agenda of outlining message effectiveness, contemporary theories of persuasion (e.g., Eagly and Chaiken 1993; Kruglanski et al. 2004; Petty and Cacioppo 1984) regularly make use of 'strong' and 'weak' arguments to bring about attitude change. That some arguments are defined as stronger than others immediately raises the question of why this should be so—and suggests that normative questions about argument strength are also central to the study of persuasion (Areni and Lutz 1988; O'Keefe and Jackson 1995; Petty and Wegener 1991).

The critical role of norms in the field of argumentation is representative of their wider popularity: Norms have been central to the study of human cognition whenever people have asked questions about rationality (Evans and Over 1996; Nickerson 2007; Nisbett and Ross 1980; Oaksford and Chater 1998). Cognitive programmes such as that initiated by Tversky and Kahneman investigating both the calibration and underlying mechanisms of people's 'intuitive statistics' have placed a huge emphasis on normative questions about behaviour. The seemingly considerable shortfalls in rationality exhibited by people's judgments that are apparent in phenomena such as the conjunction fallacy (Tversky and Kahneman 1982, 1983) are topics of continuing interest to this day. Research has focussed on elucidating particular 'heuristics' and 'biases', but has also posed broad questions about the extent to which people are Bayesian—or 'rational'.

In the equally sprawling literature on decision-making and rational choice, the prescriptive rules of decision theory (in both its objective and subjective form) have guided the evaluation of decision-making behaviour in economic and psychological experiments (see, e.g., Edwards and Tversky 1967; Pratt et al. 1995). Similarly, much of the literature on human reasoning has focussed on logic, and human deviations from it. Studies of the Wason selection task (Wason 1968), syllogistic reasoning (e.g., Johnson-Laird and Bara 1984), and reasoning with conditionals (Evans and Over 2004; Evans et al. 2005; Johnson-Laird and Byrne 2002; Manktelow and Over 1991; Oaksford and Chater 2003) are ubiquitous in cognitive psychology—and usually motivated by a desire to document people's ability to reason according to putative normative standards. Finally, social psychologists of the 1960s (e.g., McGuire 1960) and 1970s (e.g., Wyer and Goldberg 1970) used logical and probabilistic norms to evaluate the consistency of beliefs, or measure belief change (e.g., Edwards 1961; Slovic and Lichtenstein 1971).

In these diverse studies of human thought the role of norms has been twofold. First, the extent to which human behaviour matches up to these putative 'gold standards', and therefore the extent to which we might rightly claim to be rational, is of fundamental interest in its own right. It is also the question that has dominated the reception of this work in areas beyond psychology. Second, specific deviations have been critical in formulating and testing actual process theories of how humans go about these tasks (e.g., Kahneman and Tversky's Prospect Theory, 1979, or Johnson-Laird's mental models theory in the domain of logical reasoning, 1983).

At the same time, deviations from these supposedly rational standards have led to discussion about the standards themselves. In particular, spearheaded by Simon's notion of 'bounded rationality' (Simon 1982), researchers have come to focus on the adaptive value of cognitive strategies as a normative standard (e.g., Gigerenzer 1991; Gigerenzer and Selten 2002; Gigerenzer and Todd 1999). On the one hand, this has led to the method of 'rational analysis' as a means for studying cognition (see Anderson 1990; Chater and Oaksford 2008; Oaksford and Chater 1998, 2007 for overviews). Here, an optimal computational solution to an 'environmental' problem faced by an organism is identified, and provides a functional explanation of the organism's actual behaviour which is viewed as an approximation of that strategy. This framework has now been applied far beyond the reaches of judgment, reasoning, or decision-making, thus broadening the issue of rationality to novel domains such as memory (Anderson and Schooler 1991) or categorization (Anderson 1990). On the other hand, the emphasis on adaptive value has also led some to question whether probability, logic and decision theory provide the appropriate normative standards for rationality at all (for discussion see, e.g., Kahneman and Tversky 1979; Noveck and Sperber 2004; Stich 1990).

It seems fair to say that the rationality debate is alive and well after nearly half a century of empirical evidence. Whether the right standards are being invoked in experiments, whether people adhere to these rules, whether people *should* be adhering to them, and what people's behaviour in experiments tells us about human rationality are all questions that are still up for grabs. But conducting research into people's ability to respond in line with putatively rational norms presupposes something very important—that it is *possible* to derive norms for human judgment at all. What are norms, and how might they be justified?

Given how central normative standards have been to inquiry about human thought, it is surprising how little discussion there has been (in psychology at least) of *why* a rule or standard should be considered normative. Typically, the status of putative norms is simply assumed (taken to have been settled by work in other areas, such as philosophy) or, on those occasions when it is questioned, it is simply denied outright (see, e.g., Bishop and Trout 2005). The avoidance of any detailed discussion about the nature of normativity by psychologists, however, seems problematic in light of the fact that there has been extensive debate about which, of several possible candidates, might be the appropriate norm for a given task. For example, the literature on logical reasoning has now witnessed almost 15 years of debate about whether people's reasoning with conditionals should be measured against the standards of classical logic, some non-classical logical variant or the standards of probability theory (e.g., Oaksford and Chater 1994, 2003; Evans and Over 2004). Similarly, proponents of adaptive heuristics have, in

many contexts, disputed that probability theory and decision theory characterize a behavioural ideal (Gigerenzer and Todd 1999).

It is difficult to settle questions about what should be considered the appropriate norm for a particular behaviour without an understanding of *where* normativity derives from in the first place. At the same time, there have been even more radical criticisms of the normative focus of cognitive research. Some have gone so far as to suggest that it is impossible to demonstrate human irrationality using the experimental method at all (Cohen 1981), or that norms should be abandoned in the study of reasoning altogether (Elqayam and Evans 2011). It seems important that these debates involve an understanding of how and why something might be considered a norm for behavior in the first place.

In this paper, we take a step back from *empirical* debates about human rationality, and consider the question of how norms might be identified in general**.** We introduce insights from discussions of normativity found in legal philosophy, and epistemology. Law as a body of norms, and theoretical insights from legal philosophy, turn out to be a very useful source here. We then consider the kinds of foundations for norms that legal philosophers and epistemologists have posited, as a means of granting normative status in the realm of reasoning**.** Using these general considerations, we examine the normativity debate in argumentation, and consider evidence from the Bayesian and pragma-dialectical theories of argumentation in relation to their philosophical positions on normativity. In conclusion, we propose firstly that it is possible and desirable to invoke normative standards for rational argumentation, and secondly that a Bayesian approach provides solid principles with which to do so.

2 The normative question

What makes something normative? This question, or variants of it, can be found in extremely diverse fields of enquiry such as ethics, epistemology and legal philosophy, and capturing the notion of normativity is not a straightforward task. A starting point is the etymology of the word "norm", which derives directly from the Latin *norma*, the term used to describe a builder's square (Railton 2000). The purpose of a builder's square is to allow *actual* cuts to be compared to an objective standard of correctness (i.e. a geometric right angle). When deviations are noted between the *actual* cut and the *norma*, corrections are made to the cut rather than the tool. The tool provides a normative standard with which to evaluate the cut, and epistemological, moral, or legal norms provide standards with which to evaluate human behaviour. However, the analogy really only captures some of what makes normativity such a controversial philosophical topic, particularly when applied to human rationality. What if, instead of a straight cut, one wished to evaluate a curve? Now, the tool seems inappropriate, and doesn't seem to provide normative guidance at all. In doubt is the applicability of the norm, which demonstrates that norms must somehow be derived—and that it is possible to derive the wrong ones.

In fact, philosophers have recognised for a long time that simply observing whether a particular behaviour matches some prescriptive standard is not all that is required to understand the concept of normativity (see e.g., Finlay 2010). If it were, then there

would be apparently normative behaviour all around us. Using the correct side of the staircase in a busy train station is an example of something that we *ought* to do. If mammals do not lay eggs, and horses are mammals, then it *ought* also to be the case that horses do not lay eggs. But somehow, neither of these candidate conceptions of normativity seems to capture the essence of what it means for something to be normative.

On the one hand, conventions for navigating the stairwells of train stations are fairly arbitrary—and would seem to confer compliance rather than normativity. If normativity in rational argumentation could be equated with a shared set of agreed procedural rules, then "being rational is like a musician being in tune…all that matters is that we reason harmoniously with our fellows" (Chater and Oaksford 2000). On the other hand, that horses do not lay eggs necessarily follows from the fact that mammals do not lay eggs, and this also seems to eliminate the need to label the statement normative. To have a distinctive place in the world, normativity cannot simply be the *must* of conceptual necessity (Railton 2000).

2.1 Epistemology and normativity

Assuming that it is possible to identify a definition of normativity that is neither post hoc and arbitrary nor simply subsumed by necessity, where might this notion of normativity come from—that is, how can norms be derived at all?[1] Epistemologists have struggled for centuries with the question of what makes something normative, and it is still disputed to this day. Indeed, a not uncommon view is that *any* attempt to provide normative justification for beliefs is doomed to failure by one of three routes (known as the 'Munchhausen Trilemma'—Siegel and Biro 2008): (1) invoking an infinite epistemic regress (whereby A is justified by B, which is justified by C…etc); (2) developing a viciously circular series of arguments that depend on each other for their validity (e.g., normative beliefs are those that are justified—and justified beliefs are normative); or (3) introducing an arbitrary point at which a belief is simply declared 'justified', and the search for further justification aborted.[2]

Infinite epistemic regress is akin to a persistent child, who refuses to stop asking the question 'why?' Yet this is actually one of the hardest questions in epistemology—at what point does something become *just so*? Some things do seem to possess this property of self-evidence—analytic truths, for example, which are simply true by definition. But epistemologists have been cautious in granting the concept of self-evidence legitimacy beyond a selective group of logical and mathematical principles—and even

[1] Our focus here is not on what norms might be (see Finlay 2010 for an overview), but how they might be derived. The two questions are closely related, but the material presented here arguably remains relevant on all views of normativity itself. For example, on a neo-Aristotelian view such as Thomson (2008), 'rational argument' might be viewed as a normative kind, and the norms for argumentation we discuss are ways of spelling out what makes such an argument 'better' or 'worse', and why.

[2] It is worth noting, that Atkinson and Peijnenburg (2009) have recently demonstrated that from a probabilistic perspective, an infinite regress does not preclude a definite degree of support being bestowed on a claim. Likewise, not all circular dependencies are vicious from a probabilistic perspective (Atkinson and Peijnenburg 2010; Hahn 2011b). In light of this, the popular argument from regress or circularity can no longer be assumed to hold without closer consideration of the specific case at hand.

these are not uncontroversial (e.g., Finnis 1980). So there is an important and on-going debate about whether and how beliefs can be justified (Audi 2002; Siegel and Biro 2008), but it is possible to identify two broad approaches.

The first is known as *foundationalism*. This holds that if we know anything at all, we must have at some point derived 'direct' knowledge (often held to be sensory information, although it is well known that sensory knowledge is 'constructed' as much as it is 'perceived'). Foundationalists suggest that our knowledge and beliefs must be anchored in something more concrete—some kind of directly observed evidence. So, our knowledge about when to cross the road can be traced back to sensory knowledge about speed, depth and colour cue perception. The opposing view is *anti-foundationalism* (or *coherentism*)—whereby beliefs are justified if they are coherent with the other beliefs that an individual holds. On this view, holding a justified belief is more like answering a question in the light of relevant information than deducing a theorem from a set of axioms (Audi (2002, p. 196)).

Clearly the issue of normativity and justification in epistemology is a fundamental one, and epistemology offers a rich source of information from which to draw insights about normativity in argumentation. However, the epistemological approach to normativity tends to operate at a fairly high level of abstraction (typically concerning beliefs and knowledge), whereas argumentation and reasoning theorists seek to assess *specific* norms, and how they might be justified as principles of rational debate. Because of this discrepancy in analytical approach, we have also used legal philosophy to guide our analysis of normativity.[3] Legal philosophers have sought to deal not only with the wider question of whether universal norms are possible and how they might be discerned, but also the more immediate question of how an *individual* rule, contained in a specific legal system acquires normative status. This fits with our current goal of taking rules for argumentation or reasoning and considering how their normative status might be founded. It also represents a novel integration of legal theorising on normativity with psychological research on normative standards for behaviour.

2.2 Legal philosophy and normativity

Several broad strategies for bestowing normativity can be found in legal theory, which map closely on to the foundationalist/coherentist distinction that epistemologists have pursued. The first strategy seeks to derive normative status from other norms (e.g., Kelsen 1941), and is analogous to the foundationalist view in epistemology. On this account, normative power is derived from deferral to ever more fundamental and a priori truths. Many epistemologists (e.g., Railton 2000) consider this account to be highly problematic, as derivation of normativity from other norms faces the problem of a potentially infinite regress. Kelsen seeks to avoid the regress problem through the adoption of a single, otherwise content-less basic norm ("*Grundnorm")* that he claims must form the underlying basis for a legal system. His theory is an attempt

[3] Given that law, next to science, constitutes our most comprehensive system for dealing with the world, and legal philosophy provides meta-theory to it that has been developed over centuries, it seems surprising that legal philosophy is not given more prominence in epistemology and ethics, and even extensive works on normativity itself (such as Thomson 2008) make little or no reference to it.

to find a point of origin for *all* law, on which basic legal principles (and the specific laws that derive from them) obtain their legitimacy. However, short of positing some sort of a priori 'super rule' that could infuse attempts at knowledge acquisition with a normative seal of approval, many epistemologists (e.g., Railton 2000) have rejected the idea that normativity is a feature of the world in some immutable sense.

A second strategy, analogous to the epistemological approach of coherentism, seeks to derive normativity from assent or recognition (e.g., Hart 1961). However, this strategy too has faced much philosophical criticism. Deriving normativity from recognition raises the question of when, if ever, the normative can be derived from the descriptive, that is, *ought* inferred from *is* (and vice versa, see e.g., Hume 1740; for a discussion of the is-ought fallacy in the rationality debate, see Stanovich 1999). From the fact that I *ought* to be at my desk, it does not follow that I actually *am* at my desk, and from the fact that I *am* at the bar, it does not follow that I *ought* to be there**.** Breaching the ontological divide between *is* and *ought* means that norms are permitted to be normative simply because they are the norms that we follow. Blending the descriptive with the normative in this way would seem to drain normativity of its coercive power—if we can only say that one norm is as good as another, so long as it is agreed upon, then it becomes difficult to evaluate behaviour as right or wrong in any meaningful sense.

Hart (1961) responds to this criticism in two ways. First, as a legal positivist, the normativity he is defining for the legal system is not an absolute one, in the sense of immutable universals. Like anti-foundationalist epistemologists, Hart is not seeking to derive truths with a single point of origin, and law is viewed as separate from morality. Assent or recognition seeks to define only a qualified obligation for those that fall under the scope of this assent. At the same time, however, theft does not simply become legal for those who want to steal, because it is a qualified assent that counts; specifically it is the recognition of a particular group—the officials administering the system—that counts. This strategy of deferring to an 'expert' as a method of protecting normativity against accusations of arbitrariness has also been entertained in relation to norms for reasoning (Stich 1985, 1990; Stanovich 1999; Elqayam 2003), and, as we describe in the next section, theories of argumentation.

Having identified the different notions of normativity that have developed in epistemology and legal philosophy, we will now focus on two competing normative theories of argumentation—pragma-dialectical theory and the Bayesian approach—and apply the broad strategies that epistemologists and legal philosophers have identified to these candidate norms for argumentation. As shall become apparent, the issues raised here generalize widely to other aspects of human reasoning and thought.

3 Normativity and argumentation

As discussed above, standards of rational inference have been a topic of interest since antiquity. For much of this time, logic (in one form or another) has been the putative standard against which arguments are evaluated. However, there has been an increasing perception, fuelled by a wide variety of considerations, that logic cannot provide an appropriate standard by which to judge argument strength (see, e.g., Boger 2005;

Hamblin 1970; Heysse 1997; Johnson 2000; see also Oaksford and Chater 1991, 1998; Evans 2002 for critiques of logicism in the field of reasoning more generally).

This has led on the one hand to a dialectical (or rhetorical) approach to understanding argumentation (Bailenson and Rips 1996; Brem and Rips 2000; Rips 1998, 2001; Toulmin 1958; van Eemeren and Grootendorst 2004,; also Slob 2002 for discussion), based on the assumption that the problem with logic as a standard for argumentation is its inability to account for the myriad of practical influences on the acceptability of arguments in a dialectical context. To understand argumentation one needs to understand arguments 'in use', that is, one needs to appreciate that individual arguments are embedded in a dialectical sequence of claims and counterclaims. This dialectical approach has also been taken up by researchers concerned with practical questions about the development of critical thinking skills in children (Driver et al. 2000; Kuhn and Udell 2003; Simon et al. 2002), and the communication of science (Broome 1999; Korpan et al. 1997; Norris et al. 2003, but see also Corner and Hahn 2009).

The limitations of logic have also led to the rise of Bayesian probability as an alternative calculus for the evaluation of argument strength (e.g., Hahn and Oaksford 2007). Here, the problem with classical logic is the imposition of a binary normative standard that permits argumentation to be only valid or invalid—and nothing in between. Moreover, logical inference is fundamentally about truth preservation, rather than capturing *changes* in beliefs (Hahn and Oaksford 2007). Even more fine grained and multi-valued logics do not, therefore, get at the heart of the problem of informal argumentation: How much change in existing beliefs should new evidence bring about?

Hence, there are now two complementary sets of purportedly normative theories of argumentation: *procedural* theories that propose rules for dialogical exchange, such as pragma-dialectical theory (e.g., van Eemeren and Grootendorst 2004; but also Alexy 1989) and attempts to establish an epistemic framework for the evaluation of argument *content*, such as Bayesian theory (see Hahn and Oaksford 2007, but also Korb 2004, and Goldman 2003). These two types of normativity in argumentation map well onto the legal and epistemological typology identified above—pragma-dialectical theory attempts to derive norms from assent or recognition (i.e. coherentism), while the Bayesian approach seeks to ground standards of rational argument in axiomatic mathematical principles (i.e. foundationalism). Normative frameworks for informal argument consequently provide a rich testing ground for examining questions about normativity.

3.1 Pragma-dialectical normativity

van Eemeren and Grootendorst (2004) claim that argumentation must be seen as a social act, designed to resolve a difference of opinion. They emphasise the idea of an idealised model of critical discussion, a method by which speech acts can be critically evaluated in an argumentative discourse. Van Eemeren and Grootendorst propose that a series of social norms and unwritten rules (closely related to linguistic theories of conversational competence, e.g., Grice 1975) define the boundaries of argumentative acceptability. For example, a discussant who has called the standpoint of another dis-

cussant into question in the 'confrontation stage' of a discussion is always entitled to challenge the discussant to defend this standpoint (van Eemeren and Grootendorst (2004, p. 137)).

These procedural rules of conversation are analogous to Hart's (1961) strategy of deriving legal norms from prominent existing legal conventions. From the consideration of the legal example, several questions immediately arise; whose assent is relevant here, what kind of normativity is granted, and to whom does it apply?Social conventions have developed historically, and it is not clear that our pragmatic rules for discourse are any more immutable than our conventions governing clothing or politeness. In other words, the normative status that pragma-dialectical rules possess cannot simply be assumed to be of a universal nature. As in the assent-based approaches to legal rules, however, such a lack of universality does not mean that assent *cannot* bestow normativity on conventions as we currently find them (though it does raise awkward questions regarding the conditions under which these rules may change, analogous, again, to the legal situation).

Anti-foundationalist (or coherentist) epistemologists maintain that epistemological norms are not derived from a priori sources, but rather develop indefinitely, in the same way that the Kuhnian notion of paradigmatic science does (Kuhn 1970). Similarly to Kuhnian science, however, the fact that the norms are subject to temporal change does not invalidate their normativity, or bestow an unacceptable degree of relativism. Anti-foundationalists argue that as the quest to understand the natural world will never be fully completed, even our 'best' epistemological norms will ultimately be replaced someday. But equally, "since there is not any question of transcending the situation we are in at any time, there is no perspective from which we can regard them as only relativistically valid" (Knowles (2003, p. 67)). For anti-foundationalist epistemologists, the paradox of deriving norms from assent is less problematic in reality than it is in theory—epistemological norms may be mutable, but on this account they are not conceived of as arbitrary, or weak and relativistic.

It is unclear, however, whether a similar case can be made for procedural theories of argumentation. Here, we find a similar conception of normativity, but it is not at all obvious that an individual (or society) is incapable of 'transcending' a conversational context. In fact, it is straightforward to propose alternative procedural rules that, although unfamiliar, could plausibly have developed through a process of assent. For example, what if there was a rule requiring that in disputes involving more than two people, an individual could only respond to the person who spoke immediately before them? Of course, it is possible to think of disadvantages to this rule, but it is equally possible to construct a case in its favour: the rule prevents confusion, promotes orderly conduct, guards against two points being discussed at the same time, etc. While it is almost impossible to transcend epistemological principles, the possibility of conceiving of worlds where the procedural rules of argumentation are radically different suggests that the stance of anti-foundationalists in relation to mutable norms is not really tenable for proponents of pragma-dialectical theory.

Perhaps more troubling for pragma-dialectical theories, though, is the *type* of assent that pragma-dialectical theory invokes. This is because there exists, in our view, an important distinction between what we will call *developmental* assent and *evaluative* assent. The pragma-dialectical conception of normativity expressed in the

idea of the critical discussion is that normative models of argumentation are simply idealised expressions of individuated behaviour that have accumulated by *developmental* assent to become norms. This accumulation of norms is captured in the notion of an ideally rational 'reasonable critic' (who has internalised these accumulated norms, and can subsequently ensure that ideally rational rules of debate are respected), and by the development of lists of questions that this rational critic could use to distinguish 'acceptable' from 'unacceptable' arguments.

The normative status of the pragma-dialectical approach is typically asserted, rather than derived, by its proponents (e.g., Hoeken 2001a,b; O'Keefe 2005), and for pragma-dialectical theories of argumentation, the situation is less straightforward than in law—in the legal example, one at least has some evidence that officials actually apply the rules in question. The rules identified by pragma-dialectical theory, are, at best, implicit in our day-to-day discourse. It cannot, therefore, simply be assumed that the right rules have been identified, or, by consequence, that they are assented to in daily practice. What is really required is *evaluative assent*; that is, would most people agree that pragma-dialectical norms *should be followed*?

To date, there has been very little direct empirical assessment of peoples' understanding and agreement to these putative rules (for exceptions see Bailenson and Rips 1996; Rips 1998, 2001; Schreier et al. 1995; Christmann et al. 2000; Mischo 2003; van Eeemeren et al. 2009; van Eemeren et al. 2012), although there is some indirect evidence that people find arguments that observe simple dialectical principles such as clarity and explicitness to be more compelling (O'Keefe 1997a,b). But should these rules be viewed as prescriptive for all members of a community within which they dominate, or are they binding only to those who directly subscribe to them? In a framework such as Hart's, individuals cannot simply opt out, because it is only the recognition of a particular group, backed by authority and sanction, that bestows normativity.

Some philosophers have suggested that recognition based normativity can be validated by deference to an 'expert source'. Stich (1985), for example, has suggested that rules of inference and epistemological principles may be justified by the process of *reflective equilibrium* carried out by a suitably expert source. Reflective equilibrium is simply the consideration (or evaluation) of an inductive process. What Stich's proposal amounts to is the suggestion that if an individual wished to establish whether a particular inductive principle was, in fact, normative, this could be derived from an assent based process so long as the assent comes from the reflective equilibrium of the individual's 'cognitive betters'—the experts in the particular inferential domain. However, Stich himself notes that it is rarely obvious who the 'cognitive betters' in any given situation are, and it is not clear who the privileged group might be in the context of procedural theories of argumentation (although see Stanovich (1999), who suggests using measures of intellectual competence).

van Eemeren and Grootendorst's (2004) reference to a 'reasonable critic', as a normative notion itself, somewhat begs the question of how this reasonable critic came to have normative authority—the locus of normativity is simply shifted elsewhere. In addition, empirical attempts to invoke pragma-dialectical criteria as a normative model of argumentation often require participants to be trained at length with complex evaluative criteria before they can perform the task (e.g., Hoeken 2001a,b). Bearing in

mind that the notion of a 'reasonable critic' has been developed using an assent based process, this seems a unreasonable expectation of what a reasonable critic should be capable of. The notion of normativity in the assent to social conventions remains, therefore, elusive.

3.2 Bayesian normativity

By contrast, as an example of an epistemic theory of argumentation stands the recent Bayesian conception of argument strength. This account has been developed to provide a formal treatment of a range of classic *argument fallacies* such as the argument from ignorance, circular arguments or slippery slope arguments (Hahn and Oaksford 2007; Korb 2004; Oaksford and Hahn 2004). Such a formal treatment has been a longstanding goal in fallacy research (Hamblin 1970), and, by virtue of providing an explanation of when particular arguments are weak, the account necessarily also provides an account of when arguments are strong.

On the Bayesian account of argument strength individual arguments are composed of a claim and evidence in support of that claim. Both claim 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 claim in light of the evidence. Hence, argument strength 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. In addition to theoretical analysis, there is also experimental work suggesting that people share core intuitions derived from the account (Corner and Hahn 2009; Corner et al. 2011; Hahn and Oaksford 2007; Harris et al. 2012; Oaksford and Hahn 2004).

To be clear, epistemic norms and procedural rules are categorically different—procedural standards in argumentation may be in place to ensure that epistemic rules are being observed (e.g., a procedural 'norm' for clarity in reasoning style might help ensure that reasoners' beliefs are updated rationally—a prerequisite for Bayesian updating is that the evidence is clearly presented). But as procedural theories of argumentation are frequently treated as if they provide normative guidance in and of themselves, a better understanding of the differences between the normative status of procedural and epistemic theories of argumentation seems a desirable goal.

Bayesian probability theory shares conceptual ground with foundationalist epistemological approaches, and Kelsen's (1941) proposal that there are fundamental norms which bestow normative power independently of whether they are followed or not. Regarding Bayesian theory, several potential sources of normative power need consideration. The first possibility is self-evidence.

In the context of epistemology, such a possibility has been voiced explicitly by Knowles (2003). Knowles claimed that the only cogent response to the problem of an infinite epistemic regress is to maintain that norms at some fundamental level must be "self evident, indubitable, self demonstrating or something of that ilk" (Knowles

(2003, p. 15)). Elaborating on this position, Knowles suggested that certain logical or mathematical principles might be good candidates for norms that are self-evident.

Bayes' Theorem follows directly from the axioms of probability theory—indeed, Bayes' rule is a *consequence* of the Kolmogorov probability axioms (Korb and Nicholson 2004; Schum 1994). These axiomatic mathematical statements are extremely minimal, and provide only the most elementary normative guidance. They define probabilities as non-negative numbers between 0 and 1, state that the probability of a certain event must be 1, and stipulate that the union of any mutually exclusive events is equal to the sum of their probabilities. From these three axioms and the definition of conditional probability Bayes' Theorem directly follows (see Howson and Urbach 1996). Some statisticians have gone so far as to suggest that probability theory is the 'inevitable' method of describing uncertainty, in that a very minimal set of assumptions will guarantee that any formalism for dealing with uncertainty will, in fact, be equivalent to probability theory. In other words, its basis is so simple and compelling that it is difficult to find alternatives that are genuinely distinct. This makes it an attractive method for deriving normative standards in theories of rational argumentation, reasoning and belief revision (Lindley 1982).

However, such an appeal to self-evidence might be perceived to be a cheat—especially given the contentious nature of self-evidence among philosophers. From probability theory's normative status as a mathematical object, it does not follow that its *application* to day-to-day inference is normative also. And that this application is not self-evident can be read off from the variety of alternative calculi that have been proposed to this effect (e.g., Dempster-Shafer theory, see Howson and Urbach 1996 for critical discussion), and from the debate surrounding the Bayesian interpretation of probabilities as subjective degrees of belief.

4 The Dutch Book Argument

Perhaps the most famous argument in the literature on Bayesian rationality is the Dutch Book Argument (DBA). The DBA has served as the central normative justification for Bayesian theory since Ramsey first proposed it (1931; see also de Finetti 1974). It is widely referred to in the psychological literature on Bayesian reasoning (see, e.g., Oaksford and Chater 1998, 2007) but it has attracted a good deal of philosophical controversy (for a review see, Hajek 2008a). It is based on linking degrees of belief to the betting preferences of a rational person—that is, a person with (hypothetical) betting preferences that conform to the probability calculus.

A Dutch Book is a combination of bets which can be shown to entail a sure loss. Like Bayes' Theorem, a Dutch Book is simply a mathematical statement, and is philosophically uncontroversial. The Dutch Book *argument* connects degrees of belief to a (theoretical) willingness to bet by assuming that a person with degree of belief X in a proposition P would be willing to pay up to £X to bet on P. Being Bayesian—that is, being in possession of degrees of belief that conform to the probability calculus—provides immunity from Dutch books (and only being Bayesian does so, see Hajek 2008a). People who have degrees of belief that do not satisfy the Kolmogorov axioms can be made to suffer a sure loss in a betting situation, as it is possible to construct

a Dutch Book against them. As accepting odds that lead to a sure loss in a betting situation would seem to be uncontroversially irrational, the DBA offers a *normative* justification for being Bayesian that is directly based on the axiomatic principles of the probability calculus, and an extremely basic set of conditions for economic rationality (i.e. never entertain odds where every outcome entails a loss for yourself). When translated into argumentation, the DBA simply ensures that reasoners do not have conflicting or inconsistent degrees of belief in a hypothesis.

The DBA does not depend on anybody actually winning or losing money—or anybody even betting at all. As several authors, including Christenesen (1996) have shown, it is the principle of being vulnerable to a sure loss that is the essence of the argument. The argument's force depends on seeing Dutch Book vulnerability not as a practical liability, but rather as an indication of an underlying inconsistency.

The simplicity and elegance of the DBA as normative justification for Bayesian theory has attracted much support (e.g., Davidson and Pargetter 1985). However, as Armendt (1993) notes, one way of judging the significance of an argument is by "the number and variety of (attempted) refutations it attracts", and by this measure the DBA is very significant indeed. As will become clear in the following section, many critiques and defences of the DBA exist. In reviewing the literaure on the DBA as normative justification for Bayesian theory, criticisms of the argument seem to fall naturally into two broad categories—those that posit caveats to its universality, and those that question whether the DBA does enough to justify its position as the normative determinant of Bayesian rationality. In what follows, we have used this typology to structure our examination of DBA critiques.

4.1 Criticisms of the DBA—type I

The first category of criticisms of the DBA are those that question whether the simplicity of the argument—that the reason one's beliefs should conform to the probability calculus is because this provides immunity from betting losses—really captures the range of situations we might encounter where belief coherence is important. Waidacher (1997), for example, argues that the DBA does not provide a normative foundation for theories of rationality because it only applies to situations with a particular formal structure—specifically, where there is a linear relation between degree of belief and payoffs. Unless we accept the "far-reaching and highly implausible hypothesis" that all the situations we face in our life can be faithfully modelled by this hypothetical structure, then the DBA provides normative justification only for a limited range of situations, and consequently is not enough to provide the basis for a normative theory of rationality. Similarly,Davidson and Pargetter (1985) note that the DBA is based on the assumption that all parties have equal access to knowledge about the outcomes of bets—whereas in reality, inequalities in informational access may exist.

There is a sense in which arguments such as these are ultimately only capable of adding caveats to the DBA, as they do not evaluate the DBA as a source of normative authority, or the link it provides between betting preferences and degrees of belief. Undoubtedly, it is *possible* to conceive of situations where one might wish to accept a Dutch Book, or where betting preferences and degrees of belief are not

systematically related. Perhaps you are eager to impress a new acquaintance, and consider the financial losses you incur in irrationally accepting bets with a guaranteed loss to be a small price to pay for their happiness. But are these cases typical, or merely the exception that proves the rule? Waidacher's (1997) argument amounts to an objection about assuming the value preferences of agents, which Sibler (1999) dismisses as "misidentifying relatively superficial problems in the application of utility theory as potentially devastating flaws in its foundation" (Sibler (1999, p. 249)).

Arguments such as these can be dealt with by adding a *ceteris parabis* clause to the DBA—all other things being equal, it is rational for your degrees of belief to obey the axioms of the probability calculus. Other authors have suggested that there are situations where betting preferences and degrees of belief may diverge, and that therefore the DBA does not do enough to justify its position as the source of normative authority for being Bayesian. Kennedy and Chihara (1979) suggest that playing intentionally poor hands in a game of Poker (i.e. knowingly allowing Dutch Books to be made against you) may be a rational strategy in the long run, as it may convince your opponent that you are a less sophisticated player than you are. Having established a false sense of security in your opponent, by losing a series of small bets, you may then stand a better chance of winning a big pot of money later on (the well known technique of 'hustling').

An economic analogue of this betting strategy is the practice (commonly employed by large retailers) of running certain product lines at a loss. Having enticed customers in by selling some products at an unprofitable rate, they are more likely to sell products on which they are making a profitable return. Kennedy and Chihara claim, therefore, that there may be situations in which it is rational to accept Dutch Books.

However, in our view Kennedy and Chihara fail to identify the crucial feature of these 'long run' strategies: The only reason that small losses (i.e. minor violations of the probability calculus) can be permitted in the short term, is that larger profits (i.e. better than 'fair' bets) are ultimately achieved. Large retailers can only 'loss lead' on certain product lines because it is good for their business overall. The hidden assumption in Kennedy and Chihara's argument is that loss making strategies are *only* rational because at some point the pendulum will swing the other way, and your poker opponent (or the consumer) will be persuaded to accept worse losses than the ones you suffered. Local losses must ultimately be counterbalanced by global gains, or else the acceptance of Dutch Books can no longer be claimed to be a rational strategy.

It is not enough, therefore, to demonstrate that people might sometimes prefer to maximize other utilities. But it is also insufficient to point to evidence that people pursue what appear to be non-normative strategies. This is because norm or value conflict does not negate normativity. This is readily apparent in law, where rules are not without exception. The killing of another individual is prohibited and sanctioned in British law, yet there are several 'full defences' against a charge of manslaughter, such as using reasonable self-defence against an attacker. Despite these exceptions, the norm clearly remains. In doubting the normative status of the rules governing manslaughter under British law one would have to show not just that there are exceptions to the

rule, but that the normative power of the rule was consistently challenged.[4] The same level of refutation is required for normative theories of rational argumentation, and it is not clear that such a refutation can be formulated against the DBA as normative justification for Bayesian rationality.

4.2 Criticisms of the DBA—type II

The second type of criticism of the DBA is more substantive—that coherence with the axioms of probability is not in fact a necessary (or sufficient—see Rowbottom 2007) condition of rationality. For example, Hajek (2005) has claimed that proponents of the DBA have ignored the logical compliment of the argument—that when probabilistic coherence is violated, you are equally as likely to accept a 'Good Book' as a Dutch Book (with a 'Good Book' being a set of betting preferences that guarantee you a sure win). Given that no-one would argue that accepting a sure win is irrational, how can probabilistic coherence be a necessary condition of rationality?

Adding the extra assumption that we are more likely to encounter 'Dutch Bookies' than 'Good Bookies' (i.e. that people are more likely to take advantage of probabilistic incoherence than reward it) seems unacceptable. As Hajek (2005) has observed, susceptibility to a Dutch Book is something that is independent of what sorts of other people happen to be around—in fact, there need not be any other people around at all. Hajek provides an answer to his own puzzle, however, by proposing that the traditional DBA should be modified, such that instead of positing that it is rational to accept fair, and *only* fair, betting quotients, the DBA should state that it is rational to accept fair or *better than fair* (i.e. favourable) betting quotients. Stated in this way, the DBA ensures that only sure-loss violations of probabilistic coherence are irrational, and the normative power of the argument is restored.

Several authors (see below) have proposed, however, that it may still be irrational to insist on adhering to criteria of probabilistic coherence. Sibler (1999) has suggested that probabilistic coherence is not consistent with 'instrumental' rationality (on instrumental rationality more generally see, Broome 1999; Finlay 2006; Schroeder 2009 but also Kelly 2003; Kolodny 2005), in the sense that attempting to become coherent merely because of the logical possibility of becoming the victim of a Dutch Book would involve such extensive cognitive effort, that it might prove instrumentally irrational itself. Sibler's argument echoes strongly the claims made by proponents of the 'bounded rationality' approach to human reasoning, which we discuss in detail below. What matters in the present context, though, is not whether people actually *are* Bayesian, but whether they *think they ought* to be. In other words, what is critical here once again is *evaluative* assent. This is measured not by the number of people observing the norm, but by the number of people agreeing that the norm is, in fact, a good one. Behavioural data cannot impact on a norm's integrity per se (a similar claim has been made by Cohen 1981, in defence of human rationality in general); rather actual

[4] The treatment of norm conflict in law and legal philosophy is arguably more subtle than treatment of the issue within meta-ethics and debate on normativity (see e.g., Thomson 2008; Finlay 2010), and an area where closer consideration of the legal example would seem beneficial.

behaviour is at best, a weak indicator of evaluative assent. We are not aware, however, of any studies measuring peoples' acceptance of probabilistic consistency as an ideal (although see Slovic and Tversky 1974, for evidence that educating people about the axioms of rational choice does not necessarily encourage them to use them).

Hookway (1993) has claimed that there is no necessary link between the normative status of a principle and our adherence to it. For example, a group may unanimously agree that being honest or open-minded is a positive trait (and a norm to be followed), but still fail to be honest or open-minded. In fact, it is hard to imagine that people would not, in general, accept that consistency and the immunity from Dutch Books that the probability calculus conveys are minimal standards of rationality that they would *like* to comply with. So the basis for the normative authority of the DBA is a composite one that mixes both derivation and assent. Assent underwrites the rationality of maintaining consistent degrees of belief (or conversely, avoiding Dutch books). Derivation from this then grants normative status to Bayesian probability as a means of achieving this goal or standard. Again, this raises the question of who assents, and what evidence there is for this assent. The simplest answer is that people agree for themselves; rationality is a matter of individual choice, and people are free to pursue irrational strategies if they so choose. All the DBA proposes is that a rational person should exclude the possibility of negative consequences (i.e. sure losses) whose unacceptability seems universally recognisable—a recommendation with appeal that it is difficult to dispute.

Thus far—for the sake of simplicity and clarity—we have only considered the *synchronic* DBA; that is, a combination of wagers offered *simultaneously*. But a different category of DBA also exists—the *diachronic* case, which considers combinations of odds offered at *different times*. Moreover, it may be argued that given the natural focus of argumentation on belief change (see e.g., Hahn and Oaksford 2007), it is particularly important to establish Bayes' rule as a dynamic norm of rationality.

Many of the same putative counter-arguments return, but now with a diachronic twist: for example, the bookable agent will not actually accept a bet, because 'looking ahead' and calculating whether there are any traps lurking in the combined odds of combinations of bets (Maher 1992) will lead to their rejection (an argument that is refuted in Skyrms 1993). Moreover, it has been argued that diachronic Dutch Books face unique challenges, because they give rise to additional circumstances where violations may occur—in particular, odds which seem fair when considered individually can actually be shown to result in a sure-loss over a longer sequence of bets. So, argue critics, if the normative basis for Bayesian updating works in the short term but not in the long term, can it in fact be rational at all? (see, e.g., Douven 1999). Thus the status of the diachronic DBA is frequently viewed as less established than the synchronic version (Hajek 2008a).

Returning to Armendt's (1993) proposal that the number and variety of refutations the DBA attracts is an indication of its importance, one final comment should be added in defence of the DBA, and its normative status. Armendt himself has questioned whether the assumption that bets in a DBA are value independent always holds (essential if the DBA is to proceed from the axioms of probability). Some (such as Bacchus et al. 1990) start out with the explicit goal of destroying the DBA, but manage only to prove that there may (although they do not specify them) be other ways

of conceiving of rational behaviour—that perhaps the DBA is not the only path to epistemic integrity (Douven 1999). But does this make the DBA unacceptable as the normative basis for Bayesian rationality?

We argue that on the evidence presented here, it does not. The indispensability of formal mathematical theories (i.e. the axioms comprising the DBA) in defining rational thought more generally is widely acknowledged, even by those who raise challenges against them. As the wide-ranging discussion of different forms of reasoning in Oaksford and Chater (2007) demonstrates, the straightforward principles of Bayesian rationality underpin many of the fundamental processes of human cognition. And Armendt provides a succinct repost to those who would prematurely abandon the DBA:

"Demands that we assume nothing and prove strong conclusions, however the demands are disguised, are unreasonable…(A)n appropriate response is to demand from the critics something better. A Bayesian's admission that his theory can be improved, seen in these terms, is not thereby an admission that the current theory is nonsense. And the fact that nobody can (correctly) prove something from nothing does not make every theory equally good or bad" (Armendt 1993, p. 20).

The DBA covers a lot of ground using an extremely minimal set of assumptions. Furthermore, while it may not perfect, it is also not the only plank in the normative justification of Bayesian rationality.

4.3 Beyond Dutch books

Several other justification strategies for the idea that rational agents should assign degrees of belief in line with the probability calculus and revise beliefs using conditionalization exist in the literature (for a review see, Hajek 2008b). Moreover, given the ever-increasing prominence of Bayesianism within epistemology, this is an active area of research that is still seeing new proposals (e.g., Leitgeb and Pettigrew 2010a,b). The most important of these are arguments based on *representation theorems*, and arguments based on *accuracy* (see e.g., Hajek 2008b; Easwaran 2011a).

Representation-theorem based arguments exploit the representation theorems for Bayesian decision theory to argue that a basic set of (putative) rationality constraints for preferences entail representation by expected utility, and thus probability and utility functions (e.g., Maher 1993). Of course, this argument runs into the difficulty that the constraints on preferences in question have seen considerable challenge regarding their own status (e.g., Anand 1993), and one might consider the normative status of decision theory to be less well-established than that of probability theory (for other problems, see also Hajek 2008b). Nevertheless, the link with decision theory and its many theoretical and practical applications itself provides an impetus for adopting probabilities as a formal framework.

In a similar spirit are demonstrations that seemingly new and different formalisms frequently turn out to be 'probabilities in disguise' (on this issue see e.g., Cox 1946; Horvitz et al. 1986; Heckerman 1986; Snow 1998, see also, Pearl 1988 and Howson and Urbach 1996 for further references; but see also, Snow 2001). And, finally, accuracy-based justifications, presently come in a number of guises (e.g., Rosenkrantz 1992;

Joyce 1998; Leitgeb and Pettigrew 2010a,b). Accuracy-based justifications typically invoke scoring rules as are used to measure the accuracy of probabilistic forecasts (e.g., in meteorology). Rosenkrantz (1992) shows that updating by Bayes' rule maximises the expected score after sampling; in other words, other updating rules will be less efficient in the sense that they will require larger samples, on average, to be as accurate. Leitgeb and Pettigrew (2010b) demonstrate that for a suitable measure of accuracy (justified in Leitgeb and Pettigrew 2010a; but see Levinstein 2012), Bayesianism follows from the simple premise that an agent ought to approximate the truth, and hence seek to minimize inaccuracy.

Again, one can argue about these further arguments for Bayesian rationality. In particular, one may query why an agent normatively ought to have the particular further goals (e.g., coherent preferences, or minimizing accuracy etc.) that form the critical assumptions for each argument. Needless to say, any of these arguments provides the basis for the instrumental rationality of Bayesianism for anyone that *does* share the goal in question, and, perhaps equally importantly, the fact that the same norms can be derived from very different starting points itself may seem indicative of the fact that there is something privileged about them.

5 Bayesian rationality and computational limitations

One general counter-argument to an instrumental basis for Bayesian rationality merits special consideration: namely the idea that the cognitive limitations of human beings imply that they should never adopt such norms. The argument raised by Sibler (1999) that maintaining probabilistic coherence in our degrees of belief implies an 'ideally rational agent' echoes the 'bounded rationality' approach to human reasoning (Simon 1982). According to the bounded rationality hypothesis, in order to completely absorb (and therefore act on) the statistics of the environment, it would be necessary to possess computational powers far in excess of the human brain. This is an argument against maintaining Bayesian principles as a theory of rational argument that goes beyond debating whether or not people's behaviour is actually Bayesian; proponents of bounded rationality claim that it is simply not practical to expect people to be *capable* of observing Bayesian norms—at least, not without assistance (Gigerenzer and Edwards 2003). Proponents of bounded rationality argue that instead, we use a series of 'fast and frugal' heuristics to approximate normative solutions.

Rationality, under this interpretation, is relative to the performance limitations of the individual and the demands of the immediate environment, and cannot be captured by a theory that proposes absolute norms. A proponent of the bounded rationality approach might suggest that in positing norms for probabilistic coherence that are *in principle* unobtainable we have simply selected the wrong ones. Instead, we should take environmental limitations into account, and settle for standards that give us a kind of contingent optimality—rationality defined not just by normative ideals but also by cognitive constraints.

There is certainly an appeal to these arguments, and in the sense that models of bounded rationality offer methods for obtaining rational outcomes that do not depend on computational powers beyond our reach, they paint a picture of rationality that

resonates with our intuitive notion of what is 'reasonable' to expect from even the most rational individual. But is this sort of rationality, no matter how 'reasonable', actually normative at all? While it may well be more sensible to calibrate your rationality to standards that are within your grasp, this does not, in and of itself, make these standards normative. This is because normativity is about obtaining the *right* answer, not simply an answer that is as close to correct as can reasonably be expected given the circumstances. If normativity could be defined in this way—as an adaptive response to whatever circumstances you may find yourself in—it would confer an undesirable level of situational specificity. Adaptivity and normativity are not equivalent: mechanisms of biological evolution are clearly adaptive, but cannot be said to be 'correct' or normative in any meaningful sense of the word.

A simple example helps to highlight the differences between 'bounded' rationality and normative rationality. Imagine you have been set a particularly difficult multiplication problem to solve in your head—say 3784 * 457. For all but the most gifted of mathematicians, this calculation is too difficult to solve accurately given reasonable time constraints. A proponent of bounded rationality might suggest, therefore, that the normative course of action, given computational limitations and environmental constraints, would be to round the numbers down to something more easily calculable, and it would be difficult to argue with the reasonableness of this suggestion. Two issues immediately arise, however, that suggest that arguing that such a solution is *normative* might be misguided.

Firstly, some people are better at computing long multiplication than others. There does not seem to be any a priori method of establishing who has normatively opted to round the numbers to 3800 and 450, and who has lazily multiplied 4000 by 500 and clearly got the wrong answer. This is a definition of normativity that gives us no normative guidance whatsoever. Secondly, getting close to the right answer, given situational constraints, does not prevent us from wanting to know what the right answer *actually* is. Adopting a definition of normativity that operates according to bounded rationality actually prevents us from making a normative judgment at all. While describing human behaviour as being rational whenever it does the best possible job given all the constraints on its operation makes intuitive sense, we cannot label this behaviour normative.

So, in the same way that behavioural data demonstrating non-compliance with Bayesian rationality does little to undermine its normative status (only its descriptive validity), the observation that it is not always practical to expect Bayesian rationality to be maintained does not undermine the claim that it is derived from normative standards. As Chater and Oaksford (2000) note, optimal models in economics, animal behaviour, or psychology rarely assume that agents are able to find perfectly optimal solutions to the problems that they face. It is widely accepted across a wide range of disciplines, therefore, that there is no contradiction in positing normative standards that are difficult to adhere to, even in principle. But should we consider things that are *impossible* to be normative?

The idea that possibility is a necessary precondition of normativity has been a feature of Western legal systems, in particular in the area of contract law (see e.g., Lando and Beale 2000) for centuries. "Impossibilium nulla est obligatio" means "there is no obligation to do impossible things", and is a fundamental principle of Roman law.

However, the question of impossibility arguably does not even arise in the present context.

Behavioural data may support the claim that people cannot *always* be Bayesian, but it certainly does not support the claim that people cannot *ever* be Bayesian.[5] Even the most ardent anti-Bayesian would not suggest that being Bayesian is impossible—although it may be demanding in many practical situations (and there is a vast literature detailing the heuristics and biases people may bring to bear on situations such as these—see, e.g., Gilovich et al. 2002). In much the same way that the multiplication problem described above may be so time consuming that few would attempt to provide an exact answer, it is certainly not impossible. Of course, one can imagine ever more complex calculations that are beyond any intellect. But even these calculations have a *correct* answer—and that answer is no less correct for the lack of a person who can accurately compute it. Similarly, Bayesian rationality is no less normative for the lack of an individual who is completely probabilistically coherent in their beliefs.

Indeed, proponents of bounded rationality have shown that given the right cognitive tuition (see, e.g., Gigerenzer and Selten 2002), or environmental tools, the difficulty of complex tasks can be greatly reduced. Presumably, the provision of probabilistic tuition has no bearing on whether a norm is normative or not. Few would want to argue that the lack of a calculator would destroy the normative authority of the mathematically correct answer in the multiplication problem described above, and by the same token, few would seek to question Bayesian rationality simply because there are practical constraints on how achievable probabilistically coherent beliefs are.

The development of Bayes' nets (see, e.g., Pearl 1988) as a tool for implementing Bayesian computations also suggests that being Bayesian may not be as difficult as it first appears. Bayes' nets illustrate the power of 'conditional independence'—the idea that probabilistic information is often relatively unaffected by changes elsewhere in a probabilistic network. This suggests that although global probabilistic consistency is daunting in principle, local consistency is considerably more manageable. This is because in reality, so much of the probabilistic information we process is *conditionally* independent. For example, you are unlikely to need to update your degree of belief in the prevalence of an obscure species of lizard in the light of recent information you received about rising levels of acidity in the Atlantic Ocean, even though this evidence potentially bears on the global ecological system. This is because lizard prevalence is likely to be conditionally independent of sea acidity, given more directly related evidence—for example, the temperature of their habitat. Within a range of habitat temperatures, the level of sea acidity is effectively 'screened off', and no probabilistic changes to your beliefs are necessary.

To undermine the normative status of Bayesian norms for argumentation, therefore, it would be necessary to demonstrate not only that someone's behaviour was

[5] A potential concern that may be raised here is that believing something is a state, not an action, and that it is impossible to make oneself have a belief (see e.g., Thomson 2008; Chrisman 2008). However, being Bayesian is about assigning coherent degrees of belief, not believing per se, and, it seems no more impossible to adjust one's own degrees of belief than it does to assign (potentially in a process of iterative refinement) plausible degrees of belief to a Bayesian Belief Network (Pearl 1988). Furthermore, while it may be impossible for people to be Bayesian about *all* their beliefs, it is entirely possible that they may be Bayesian regarding some subset of them.

not Bayesian, not only that on occasion that person would not choose to reason in a Bayesian fashion, and not only that on occasion that person would not be capable of reasoning in a Bayesian fashion, but that that person would dispute the normative claim of probabilistic consistency most of the time. Given the mal-adaptive consequences of such a stance, it seems unlikely that many such individuals can be found.

6 Revisiting *is* versus *ought*

The issue of factual, cognitive limitations and their implications for normative questions of rationality provide an example of the complex relationship between 'is' and 'ought' that pervades debates about normativity. As outlined above, inferences from 'is' to 'ought' and vice versa, have traditionally been viewed with suspicion. However, more recent work within ethics and moral philosophy, economics, and psychology, has illustrated the many close connections. As these afford additional potential routes to normative foundation for both Bayesian and procedural norms, we draw together some of that literature in the remainder of this paper.

First, philosophers concerned with the fact/value distinction (e.g., Williams 1985; Putnam 2002, on the relationship between the fact/value distinction; and illicit is-ought inference as mentioned above, but also Moore's naturalistic fallacy, see Dodd and Stern-Gillet 1995), have come to argue that 'normative versus descriptive' is better viewed as a simple distinction (that may be useful in some contexts) than as a dichotomy. In keeping with this, it has been argued that many concepts (not just moral ones) contain an evaluative component that is not readily separated from its descriptive aspects (i.e. a 'thick concept', see Williams 1985; see e.g., Thomson 2008).

A similar development has taken place within economics, which has seen a surge of interest in the relationship between the normative and the descriptive in recent decades (see e.g., Hands 2010; Harris and Freeman 2008; Mongin 2006; Sen 1987; Caplin and Schotter 2008).[6] Economists have argued about the relationship between descriptive and normative in the context of debate about the status of decision-theory and game theory (e.g., Hands 2010; Starmer 2005). This debate is closely related to the rationality concerns of norms for argumentation, in that the standard method of economics has been founded on optimisation, whereby individual agents are presumed to be rational (i.e. economic agents have stable and coherent preferences as set out by expected utility theory EUT, Von Neumann and Morgenstern 1847; but see e.g., Thaler and Mullainathan 2008; Camerer 1995) and it is the goal of economic theorizing to understand aggregate behaviours that arise from the interactions of such individuals (see e.g., Lehtinen and Kuorikoski 2007).

Though conceived primarily as a normative theory, economics has seen periods in which EUT was viewed as a descriptive theory (see e.g, Friedman and Savage 1948, 1952). Even now, given overwhelming evidence of violations of rational choice theory in both experiments and field studies (see e.g., Camerer 1995), the theories of

[6] Much of this debate concerns the relationship between economics and ethics, and the extent to which the long-held view that economics should concern itself only with statements of fact, not with ethics or morals (Robbins 1932; see also Keynes 1917; Friedman 1953; Putnam 2003; Mongin 2006) is desirable or even feasible.

aggregate behaviour arising from idealised rational agents aim to be descriptively accurate; this may be possible because certain behavioural contexts provide pressures that lead individuals to utility maximising behaviour (see e.g., Satz and Ferejohn 1994) and because the behaviour of aggregate systems may be robust to the deviations from rational choice theory real agents might display (Lehtinen and Kuorikoski 2007).

At the same time, key figures in the history of EUT such as Bernoulli, Condorcet and Jallais, have justified (or criticized) it on the basis that it captures (or does not capture) the decision rule *actually applied* by those considered 'wisest' in making uncertain choices (in an Aristotelian sense), such as businessmen or insurers (Jallais et al. 2008). That they owe their practical success (at least in part) to the application of this decision-rule, gives a reason for imitating these decision-makers in order to achieve our goals. This argumentative pattern for normativity thus combines what we have called 'evaluative' assent and instrumentalism.

The relationship between 'is' and 'ought' has been even more acute in the context of game theory which features both as a normative theory, and as a predictive or explanatory theory of what people actually do—with boundaries between normative and descriptive that can on occasion be blurry (Aumann 1985; Kadane and Larkey 1983). Because game theory aims to capture decision contexts in which players' choices are interdependent (where what constitutes an optimal strategy for one player is influenced by the strategy of the other), a predictive theory of the opponent is a necessary factor both in normative attempts to identify what players ought to do, as well as in descriptive theorizing about what actually takes place. A key development, here, has been the advent of evolutionary game theory (Maynard Smith 1982). Conceived of initially for the study of biological problems, evolutionary game theory studies populations of players, examining the relative success/failure of different player strategies, typically in their evolutionary dynamics over time. This has both led to the uncovering of systematic relationships between game theoretic solution concepts and evolutionary stable strategies and has allowed game theory to implement more realistic rationality assumptions for players (see e.g., Sugden 2001; Gruene-Yanoff and Lehtinen 2010).

This gives rise to is-ought and ought-is inferences: evolutionary 'success' has normative import. This might be because such success (empirically) underwrites key assumptions (about player rationality, or solution concepts) which themselves affect fundamentally the normative claims of game theory (see e.g., Binmore 1994; Mailath 1998; see also Sugden 2001, and Gruene-Yanoff and Lehtinen 2010, for critical discussion). Or it might be because we are inclined to view evolutionary outcomes themselves as imbued with normative interpretation. If 'evolutionary forces' (biological, social, and economic) generate pressures that lead to maximization of expected utility (Binmore 1994), for example, then we might be inclined to think this lends some force to the claim that they are 'optimal'; and, conversely, we may attribute their success to the fact that they are normatively well-founded. Normative strategies should allow one to avoid errors that might be costly and "evolution will not be kind to memes that inhibit their own replication (Binmore (1994, p. 27)).

This is directly relevant to the wider project of norms for argumentation because evolutionary game theoretic modelling has been extended far beyond the domains of the biological or traditionally economic to a wide variety of behaviours and social norms (e.g., Axelrod 1984; Binmore 1994, 1998; Skyrms 1996). Thus such an approach

might be applicable not just to epistemic norms of rationality, but also to procedural ones. Though we are unaware of any such work to date, one might attempt to show that the kinds of norms posited by pragma-dialectics emerge naturally in suitably defined populations.

What is, in effect, developmental assent within a (model) population may thus be enhanced with some form of evaluative component. Moreover, there are close conceptual links here with the instrumental imitation argument identified by Jallais et al. (2008) discussed above. Not only do some evolutionary models involve agent imitation, but the replication invoked by reproductive fitness itself has an imitative feel.

Nevertheless, evolutionary approaches offer no easy answers to questions about normativity. Setting aside the technical issue of whether there really is a compelling way to evaluate 'evolutionary success'[7] the main difficulty stems from the fact that there are significant restrictions on the extent to which evolution can be viewed as an optimizing process. Evolution, in the natural world, does not continuously make a species 'better', and traits can spread throughout a population because they enhance relative reproductive success, even though they may ultimately drive the species to extinction (see e.g., Sober 1993, for examples and discussion). Furthermore, reproductive success depends on what else is there: genes or memes do not reproduce in isolation in anything other than the unlikely case that there is only one gene (or meme) to an agent (see also Sugden 2001). Consequently, authors such as Skyrms (1996, 2009) have been careful to assert that evolutionary analysis (be it cultural or biological evolution) does not tell one what to do, but rather attempts to investigate how social conventions and norms could have evolved and what alternative norms are possible.

That is not to say, however, that evolutionary dynamics in simulated systems are uninformative with regard to normativity. Though the emergence, or even dominance, of a rule or strategy does not in and of itself make that rule optimal or desirable, evolutionary success is frequently not mere happenstance. Hence modelling can shed light on putative links between optimality and success and, where a norm or strategy can be shown to have privileged connections with something the agent/organism cares about (see e.g., Okasha 2011), this lends support to claims of instrumental rationality.

Finally, there have been close links between the normative and the descriptive within psychology. As detailed in our Introduction, there has been a long tradition of research on human rationality comparing human behaviour against normative standards and recent years have seen a wave of probabilistic, Bayesian modelling of cognitive tasks as diverse as perception, memory and language (see e.g., Chater et al. 2006; but see Jones and Love 2011; but also, Howes et al. 2009; Hahn 2011a) reflecting the explanatory project of rational analysis (Anderson 1990; Chater and Oaksford 1999, 2008). Rational analysis provides a functional explanation ('why'?) of the cognitive system by seeking to understand its workings as an approximation to the optimal solution to the task in question (independently of whether that adaptation may be result from evolution and/or learning).

[7] One difficulty here being that there may be a plethora of relevant measures (see Huttegger and Zollman 2012) in much the same way as game theory in its non-evolutionary form is faced with a variety of potential solution concepts—this, in fact, being one of the problems it was hoped evolutionary game theory might address—see Sugden (2001).

In this context, empirical evidence may be used to choose between candidate normative systems. This issue arises from the project of rational analysis because the search for interpretations that render sensible the behaviour of the system have often led to re-evaluations of seeming irrationality. One of the most well-known, and most relevant to the present context, was Oaksford and Chater's (1994) demonstration that seemingly irrational, participant behaviour in logical reasoning tasks seems appropriate from a Bayesian perspective. In other words, assuming Bayesian probability as the appropriate normative framework instead of classical logic made sense of actual behaviour. This led to a rejection of the material conditional as an appropriate formalisation of natural language conditionals (see also Evans and Over 1996) and a rehabilitation of the (logical) 'fallacies' of affirmation of the consequent and denial of the antecedent as probabilistically reasonable inference schemes (Chater and Oaksford 2000; Oaksford et al. 2000; Oaksford and Chater 2007).

Elqayam and Evans (2011) view such 'adjudication' (i.e., a choice between competing normative systems made on the basis of empirical data) as an illicit *is* to *ought* inference, and recommend that psychologists abandon comparing human behaviour to putative norms. But this seems mistaken, because it is not the normative status of classical logic or probability theory that is at issue here; both are viewed as normative and, as normative systems, they are compatible. Rather it is the mapping between formalism and task that is in question; that is, the issue at hand is about the application of norms to a particular instance. Again, the example of law is helpful here. Using someone else's credit card details might be considered to be theft or fraud; the decision on which norm of the criminal code is the more appropriate will draw on a range of facts, both about the real-world situation and the norms in question. Their normative status, however, is not at stake.

At the same time, it seems legitimate to use experimental data in support of normativity as much as it seems legitimate to use intuition. Intuition involves a descriptive fact. Should appeal to intuition (and at its most extreme: self-evidence) be barred from normative discourse as an illegitimate is-to-ought inference? It is hard to see that anything would be left if this were the case. Experiments on reasoning (or argumentation) can be viewed not only as tests of 'norm-conformity' in lay people, they serve to validate researcher intuitions which may turn out to be less compelling and less widespread than researchers themselves think (Hahn and Oaksford 2007; paralleling recent developments in moral philosophy, e.g., Nichols and Knobe 2007, and the wider movement towards 'experimental philosophy').

As detailed in the Introduction, a considerable range of everyday argument forms has now been investigated from a Bayesian perspective, although the list remains far from complete. Most of these are not logically valid, yet intuitively they can still be strong, given the right content. The fact that a Bayesian formalization seems to match fundamental intuitions about argument strength (and evidence in general, see e.g., Howson and Urbach 1996; Bovens and Hartmann 2003; Easwaran 2011b) across a broad range of argument types and circumstances itself seems normatively relevant. It resonates with Laplace's perception of the probability calculus as 'formalised common sense' (Laplace 1951)1814, and it may be viewed as providing support for the normative status of the calculus, much in the same way that Goodman viewed the justification of deductive rules of inference:

> The validity of a deduction depends upon. . .conformity to valid rules. . .. But how is the validity of rules to be determined?.. . Principles of deductive inference are justified by their conformity with accepted deductive practice. Their validity depends upon accordance with the particular deductive inferences that we actually make and sanction. If a rule yields inacceptable inferences, we drop it as invalid. Justification of general rules thus derives from judgments rejecting or accepting particular deductive inferences.
> This looks flagrantly circular. . .. But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either. (Goodman (1965, pp. 66–67), emphasis original)

Of course, probability theory will not always match every intuition (and, in fact, part of the insight to be obtained from adopting a formal framework is that some evaluations will be changed, e.g., Atkinson and Peijnenburg 2009) and it will not always match everyone's intuition. But this does not pose a problem (cf. Buckwater and Stich 2011), because these intuitions are not our only guide; rather they operate in conjunction with insights from adaptive modelling, mathematical derivation, and empirical success.

In sum, the normative and the descriptive are interwoven, providing further opportunities for grounding a Bayesian norm. By the same token, judgements about norms for argumentation (and with them norms for rational behaviour more generally) need not rest on single strand: as with a rope, the justification provided by these strands taken together exceeds their individual strength.

7 Conclusion

The normative question is a complex one, but we have tried to identify what seem to be the crucial aspects of normativity for argumentation. By drawing on ideas about normativity from legal philosophy and epistemology, we have sought to shed light on what sort of normative theory might be appropriate for argumentation, and what features such a theory might need to incorporate. We reviewed the claim to normativity of two very different types of theory of argument—procedural and epistemic—and it would seem that the Bayesian account meets many of the requirements for normativity as philosophers have sought to elucidate it.

By combining the self-evidence of the axioms of probability theory with the minimal economic rationale of the DBA, or a desire for accuracy, Bayesian inference seems to be based on solid normative principles that are not vulnerable to the problem of infinite regress. Assent is required to the extent that reasoners must agree that in general, benefiting from the protection Bayesianism provides against inconsistency is a good thing. Whilst people may often (for any number reasons) deviate from Bayesian

inference behaviourally, it seems unlikely that they would dispute the normative rationale of Bayesian principles. By contrast, the normative status of procedural rules of discourse is presently less clear.

These issues are not only of philosophical interest, given how fundamental argumentation is to science, politics and everyday life. Possessing a normative theory of argument strength is also essential to psychological research (Areni and Lutz 1988; O'Keefe and Jackson 1995; Petty and Wegener 1991; van Enschot-Van Dijk et al. 2003) and the Bayesian approach to informal argumentation seems to provide an answer to this longstanding problem. In summary, the notion of Bayesian normativity in argumentation can be retained for the same reason that normativity is retained generally in philosophy—there are many methods for deriving it.

Acknowledgments This work was partly conducted while the second author was at Cardiff University.

References

Alexy, R. (1989). *A theory of legal argumentation*. Oxford: Clarendon Press.
Anand, P. (1993). *Foundations of rational choice under risk*. Oxford: Oxford University Press.
Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum Press.
Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*(6), 396–408.
Areni, C. S., & Lutz, R. J. (1988). The role of argument quality in the elaboration likelihood model. *Advances in Consumer Research, 15*, 197–203.
Armendt, B. (1993). Dutch books, additivity and utility theory. *Philosophical Topics, 21*(1), 1–20.
Atkinson, D., & Peijnenburg, J. (2009). Justification by an infinity of conditional probabilities. *Notre Dame Journal of Formal Logic, 50*, 83–93.
Atkinson, D., & Peijnenburg, J. (2010). Justification by infinite loops. *Notre Dame Journal of Formal Logic, 51*, 407–416.
Audi, R. (2002). *Epistemology: A contemporary introduction*. London: Routledge.
Aumann, R. J. (1985). What is game theory trying to accomplish?. In K. Arrow & S. Honkapohja (Eds.), *Frontiers of economics* (pp. 28–76). Oxford: Blackwell Publishing.
Axelrod, R. (1984). *The evolution of cooperation*. London: Penguin.
Bacchus, F., Kyburg, H. E., & Thalos, M. (1990). Against conditionalization. *Synthese, 85*(3), 475–506.
Bailenson, J., & Rips, L. J. (1996). Informal reasoning and burden of proof. *Applied Cognitive Psychology, 10*, 3–16.
Binmore, K. (1994). *Game theory and the social contract, volume I: Playing fair*. Cambridge, MA: MIT Press.
Binmore, K. (1998). *Game theory and the social contact, volume II: Just playing*. Cambridge, MA: MIT Press.
Biro, J., & Siegel, H. (2006). In defense of the objective epistemic approach to argumentation. *Informal Logic, 26*(1), 91–101.
Bishop, M. A., & Trout, J. D. (2005). *Epistemology and the psychology of human judgment*. New York: Oxford University Press.
Boger, G. (2005). Subordinating truth—Is acceptability acceptable?. *Argumentation, 19*, 187–238.
Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.
Brem, S. K., & Rips, L. J. (2000). Explanation and evidence in informal argument. *Cognitive Science, 24*(4), 573–604.
Brem, S. K., Russell, J., & Weems, L. (2001). Science on the web: Student evaluations of scientific arguments. *Discourse Processes, 32*, 191–213.
Broome, J. (1999). Normative requirements. *Ratio, 12*(4), 398–419.
Buckwater, W., & Stich, S. (2011). Competence, reflective equilibrium, and dual-system theories. Commentary on Elqayam & Evans. *Behavioral and Brain Sciences, 34*, 251–252.
Camerer, C. (1995). Individual decision making. In J. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics*. Princeton, NJ: Princeton University Press.

Caplin, A., & Schotte, A. (2008). *The foundations of positive and normative economics: A handbook*. Oxford: Oxford University Press.
Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences, 3*(2), 57–65.
Chater, N., & Oaksford, M. (2000). The rational analysis of mind and behaviour. *Synthese, 122*, 93–131.
Chater, N. & Oaksford, M. (Eds.). (2008). *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford: Oxford University Press.
Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences, 10*(7), 287–291. doi:10.1016/j.tics.2006.05.007.
Chrisman, M. (2008). Ought to believe. *Journal of Philosophy, 105*, 346–370.
Christenesen, D. (1996). Dutch-book arguments depragmatized: Epistemic consistency for partial believers. *The Journal of Philosophy, 93*(9), 450–479.
Christmann, U., Mischo, C., & Groeben, N. (2000). Components of the evaluation of integrity violations in argumentative discussions: Relevant factors and their relationships. *Journal of Language and Social Psychology, 19*, 315–341.
Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated?. *Behavioural and Brain Sciences, 4*, 317–370.
Corner, A., & Hahn, U. (2009). Evaluating scientific arguments: Evidence, uncertainty & argument strength. *Journal of Experimental Psychology: Applied, 15*(3), 199–212.
Corner, A. J., Hahn, U., & Oaksford, M. (2011). The psychological mechanism of the slippery slope argument. *Journal of Memory and Language, 64*, 153–170.
Cox, R. (1946). Probability frequency and reasonable expectation. *American Journal of Physics, 14*, 1–13.
Cox, G. W. (1999). The empirical content of rational choice theory: A reply to Green and Shapiro. *Journal of Theoretical Politics, 11*(2), 147–169.
Davidson, B., & Pargetter, R. (1985). In defence of the Dutch book argument. *Canadian Journal of Philosophy, 15*(3), 405–424.
de Finetti, B. (1974). *Theory of probability*. New York: Wiley.
Dodd, J., & Stern-Gillet, S. (1995). The is/ought gap, the fact/value distinction and the naturalistic fallacy. *Dialogue, 34*(04), 727–746. doi:10.1017/S0012217300011082.
Douven, I. (1999). Inference to the best explanation made coherent. *Philosophy of Science, 66*, S424–S435.
Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 287–312.
Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace.
Easwaran, K. (2011a). Bayesianism I: Introduction and arguments in favor. *Philosophy Compass, 6*(5), 312–320. doi:10.1111/j.1747-9991.2011.00399.x.
Easwaran, K. (2011b). Bayesianism II: Applications and criticisms. *Philosophy Compass, 6*(5), 321–332. doi:10.1111/j.1747-9991.2011.00398.x.
Edwards, W. (1961). Behavioural decision theory. *Annual Review of Psychology, 12*, 473–498.
Edwards, W. & Tversky, A. (Eds.). (1967). *Decision making*. Middlesex: Penguin.
Elqayam, S. (2003). Norm, error and the structure of rationality: The case study of the knight-knave paradigm. *Semiotica, 147*, 265–289.
Elqayam, S., & Evans, J. S. B. T. (2011). Subtracting "ought" from 'is': Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences, 34*(05), 233–248. doi:10.1017/S0140525X1100001X.
Evans, J. St. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin, 128*(6), 978–996.
Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove, UK: Psychology Press.
Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
Evans, J. St. B. T., Over, D. E., & Handley, S. J. (2005). Suppositions, extensionality and conditionals: A critique of the mental model theory of Johnson-Laird & Byrne (2002). *Psychological Review, 112*, 1040–1052.
Finnis, J. (1980). *Natural law & natural right*. Oxford: Oxford University Press.
Finlay, S. (2006). The reasons that matter. *Australasian Journal of Philosophy, 84*(1), 1–20. doi:10.1080/00048400600571661.
Finlay, S. (2010). Recent Work on Normativity. *Analysis, 70*, 331–346. doi:10.1093/analys/anq002.
Friedman, M. (1953). *The methodology of positive economics*. Chicago: Chicago University Press.

Friedman, M., & Savage, L. (1948). The utility analysis of choices involving risk. *Journal of Political Economy, LVI*, 279–304.
Friedman, M., & Savage, L. (1952). The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy, LX*, 463–474.
Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond 'heuristics and biases'. *European Review of Social Psychology, 2*(1), 83–115.
Gigerenzer, G., & Edwards, A. (2003). Simple tools for understanding risks: From innumeracy to insight. *British Medical Journal, 327*(7417), 741–744.
Gigerenzer, G. & Selten, R. S. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. Cambridge, MA: MIT.
Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
Gilovich, T., Griffin, D. W., & Kahneman, D. (2002). *Heuristics & biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.
Goldman, A. I. (2003). An epistemological approach to argumentation. *Informal Logic, 23*, 51–63.
Goodman, N. (1965). *Fact, fiction, and forecast*. Cambridge, MA: Bobbs-Merrill.
Grice, H. P. (1975). Logic and Conversation. In D. Davidson & G. Harman (Eds.), *The logic of grammar*. Encino, CA: Dickenson.
Gruene-Yanoff, T., & Lehtinen, A. (2010). Philosophy of game theory. In U. Maki (Ed.), *Philosophy of economics*. Volume 13 of D. Gabbay, P. Thagard, & J. Woods (Eds.), *Handbook of the philosophy of science*. Amsterdam: Elsevier.
Hahn, U. (2011a). Why rational norms are indispensible. Commentary on Elqayam and Evans. *Behavioral and Brain Sciences*, 257–258.
Hahn, U. (2011b). The problem of circularity in evidence, argument and explanation. *Perspectives on Psychological Science, 6*, 172–182.
Hahn, U., & Oaksford, M. (2006a). A Bayesian approach to informal fallacies. *Synthese, 152*(2), 207–237.
Hahn, U., & Oaksford, M. (2006b). Why a normative theory of argument strength and why might one want it to be Bayesian?. *Informal Logic, 26*, 1–24.
Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review, 114*(3), 704–732.
Hahn, U., Oaksford, M., & Corner, A. (2005). Circular arguments, begging the question and the formalization of argument strength. In *Proceedings of AMKCL05—Adaptive knowledge representation and reasoning*. Helsinki.
Hajek, A. (2005). Scotching Dutch books?. *Philosophical Perspectives, 19*, 139–151.
Hajek, A. (2008a). Dutch book arguments. In P. Anand, P. Pattanaik, & C. Puppe (Eds.), *The handbook of rational and social choice* (pp. 173–196). Oxford: Oxford University Press.
Hajek, A. (2008b). Arguments for-or against-probabilism?. *The British Journal for the Philosophy of Science, 59*(4), 793–819. doi:10.1093/bjps/axn045.
Hamblin, C. L. (1970). *Fallacies*. London: Methuen.
Hands, D. W. (2010). The positive-normative dichotomy and economics. In U. Maki (Ed.), *Philosophy of economics*. Volume 13 of D. Gabbay, P. Thagard, & J. Woods (Eds.), *Handbook of the philosophy of science*. Amsterdam: Elsevier.
Harris, A. J. L., Hsu, A. S., & Madsen, J. K. (2012). Because Hitler did it! Quantitative tests of Bayesian argumentation using ad hominem. *Thinking & Reasoning, 18*(3), 311–343. doi:10.1080/13546783.2012.670753.
Harris, J. D., & Freeman, R. E. (2008). The impossibility of the separation thesis. *Business Ethics Quarterly, 18*(4), 541–548.
Hart, H. L. A. (1961). *The concept of law*. Oxford: Oxford University Press.
Heckerman, D. (1986). A rational measure of confirmation. In L. N. Kanal & J. F. Lemmer (Eds.), *Uncertainty in artificial intelligence 2*. Amsterdam: North Holland.
Heysse, T. (1997). Why logic doesn't matter in the (philosophical) study of argumentation. *Argumentation, 11*, 211–224.
Hoeken, H. (2001a). Convincing Citizens The Role of Argument Quality. In D. Janssen & R. Neutelings (Eds.), *Reading and writing public documents*. Amsterdam: Benjamins.
Hoeken, H. (2001b). Anecdotal, statistical and causal evidence: Their perceived and actual persuasiveness. *Argumentation, 15*, 425–437.

Hookway, C. (1993). *Epistemic norms and theoretical deliberation*. In J. Dancy (Ed.). (2000). *Normativity*. Oxford: Blackwell.

Horvitz, E. J., Heckerman, D., & Langlotz, C. P. (1986). A framework for comparing alternative formalisms for plausible reasoning. In *Proceedings of the 5th National Conference on AI (AAAI-1986)*, pp. 210–214.

Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review, 116*(4), 717–751.

Howson, C., & Urbach, P. (1996). *Scientific reasoning: The Bayesian approach*. Chicago: Open Court.

Hume, D. (1740). *A treatise of human nature (1967 edition)*. Oxford: Oxford University Press.

Huttegger, S. M., & Zollman, K. J. S. (2012). The limits of ESS methodology. In S. Okasha & K. Binmore (Eds.), *Evolution and rationality: Decisions, co-operation and strategic behaviour*. Cambridge: Cambridge University Press.

Jallais, S., Pradier, P. C., & Teira, D. (2008). Facts, norms and expected utility functions. *History of the Human Sciences, 21*(2), 45–62. doi:10.1177/0952695108091414.

Johnson, R. H. (2000). *Manifest rationality: A pragmatic theory of argument*. Mahwah, NJ: Hillsdale.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Johnson-Laird, P. N., & Bara, B. G. (1984). Syllogistic inference. *Cognition, 16*(1), 1–61.

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

Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences, 34*(04), 169–188. doi:10.1017/S0140525X10003134.

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science, 65*, 573–603.

Kadane, J. B., & Larkey, P. D. (1983). The confusion of is and ought in game theoretic contexts. *Management Science*, 1365–1379.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision making under risk. *Econometrica, 47*(2), 263–292.

Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research, 66*(3), 612–640.

Kelsen, H. (1941). The pure theory of law & analytical independence. *Harvard Law Review, 55*, 44–66.

Kennedy, R., & Chihara, C. (1979). The Dutch book argument: Its logical flaws, its subjective sources.. *Philosophical Studies, 36*, 19–33.

Keynes, J. N. (1917). *The scope and method of political economy* (4th ed.). London: Macmillan & Co.

Knowles, J. (2003). *Norms, naturalism and epistemology*. New York: Palgrave Macmillan.

Kolodny, N. (2005). Why be rational?. *Mind, 114*(455), 509–563. doi:10.1093/mind/fzi223.

Korb, K. B. (2004). Bayesian informal logic and fallacy. *Informal Logic, 24*, 41–70.

Korb, K. B., & Nicholson, A. E. (2004). *Bayesian artificial intelligence*. Boca Raton, FL: CRC Press.

Korpan, C. A., Bisanz, G. L., Bisanz, J., & Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education, 81*, 515–532.

Kruglanski, A. W., Fishbach, A., Erb, H. P., Pierro, A., & Mannetti, L. (2004). The Parametric Unimodel as a Theory of Persuasion. In G. Haddock & G. R. Maio (Eds.), *Contemporary perspectives on the psychology of attitudes*. New York: Psychology Press.

Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development, 74*(5), 1245–1260.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Lando, O., & Beale, H. (2000). *Principles of European contract law*. Boston: Kluwer Law International.

Laplace, P. S. (1951). A philosophical essay on probabilities trans. (Original work published 1814), F. W. Truscott & F. L. Emory. Dover.

Lehtinen, A., & Kuorikoski, J. (2007). Unrealistic assumptions in rational choice theory. *Philosophy of the Social Sciences, 37*(2), 115–138. doi:10.1177/0048393107299684.

Leitgeb, H., & Pettigrew, R. (2010a). An objective justification of Bayesianism I: Measuring inaccuracy*. *Philosophy of Science, 77*(2), 201–235.

Leitgeb, H., & Pettigrew, R. (2010b). An objective justification of Bayesianism II: The consequences of minimizing inaccuracy*. *Philosophy of Science, 77*(2), 236–272.

Levinstein, B. A. (2012). Leitgeb and Pettigrew on accuracy and updating. *Philosophy of Science, 79*(3), 413–424. doi:10.1086/666064.

Lindley, D. V. (1982). Scoring rules and the inevitability of probability. *International Statistical Review, 50*, 1–26.

Maher, P. (1992). Diachronic rationality. *Philosophy of Science, 59*, 120–41.
Maher, P. (1993). *Betting on theories*. Cambridge: Cambridge University Press.
Mailath, G. J. (1998). Do people play nash equilibrium? Lessons from evolutionary game theory. *Jounal of Economic Literature, XXXVI*, 1347–1374.
Manktelow, K. I., & Over, D. E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition, 39*(2), 85–105.
Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, NY: Cambridge University Press.
McGuire, W. J. (1960). Cognitive consistency and attitude change. *Journal of Abnormal and Social Psychology, 60*(3), 345–358.
Mischo, C. (2003). Cognitive, emotional and verbal response in unfair everyday discourse. *Journal of Language and Social Psychology, 22*(1), 119–131.
Mongin, P. (2006). Value judgments and value neutrality in economics. *Economica, 73*(290), 257–286. doi:10.1111/j.1468-0335.2006.00501.x.
Neuman, Y., Weinstock, M. P., & Glasner, A. (2006). The effect of contextual factors on the judgment of informal reasoning fallacies. *Quarterly Journal of Experimental Psychology: Human Experimental. Psychology, 59*(A), 411–425.
Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous, 41*, 663–685.
Nickerson, R. S. (2007). *Aspects of rationality: Reflections on what it means to be rational and whether we are*. Hillsdale, NJ: Psychology Press.
Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement*. Prentice, NJ: Prentice Hall.
Norris, S. P., Phillips, L. M., & Korpan, C. A. (2003). University students' interpretation of media reports of science and its relationship to background knowledge, interest and reading difficulty. *Public Understanding of Science, 12*, 123–145.
Noveck, I. A. & Sperber, D. (Eds.). (2004). *Experimental pragmatics*. New York: Palgrave Macmillan.
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. (1998). *Rationality in an uncertain world: Essays on the cognitive science of human reasoning*. Sussex: Psychology Press.
Oaksford, M., & Chater, N. (2003). Conditional probability and the cognitive science of conditional reasoning. *Mind & Language, 18*(4), 359–379.
Oaksford, M., & Chater, N. (2007). *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., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference.. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 883–899.
Okasha, S. (2011). Optimal choice in the face of risk: Decision theory mMeets evolution. *Philosophy of Science, 78*(1), 83–104. doi:10.1086/658115.
O'Keefe, D. J. (1997a). Standpoint explicitness and persuasive effect: A meta-analytic review of the effects of varying conclusion articulation in persuasive messages. *Argumentation & Advocacy, 34*, 1–12.
O'Keefe, D. J. (1997b). Justification explicitness and persuasive effect: A meta-analytic review of the effects of varying support articulation in persuasive messages. *Argumentation & Advocacy, 35*, 61–75.
O'Keefe, D. J. (2003). The Potential Conflict Between Normatively Good Argumentative Practice and Persuasive Success. In F. H. van Eemeren, J. Anthony Blair, C. A. Willard, Snoeck Francisca, & A. Henkemans (Eds.), *Anyone who has a view: Theoretical contributions to the study of argumentation*. Dordrecht: Kluwer Academic Publishers.
O'Keefe, D. J. (2005). News for argumentation from persuasion effects research: Two cheers for reasoned discourse. In C. A. Willard (Ed.), *Selected papers from the thirteenth NCA/AFA conference on argumentation* (pp. 215–221). Washington, DC: National Communication Association.
O'Keefe, D. J. (2007). Potential conflicts between normatively-responsible advocacy and successful social influence: Evidence from Persuasion research.. *Argumentation, 21*, 151–163.

O'Keefe, D. J., & Jackson, S. (1995). Argument quality and persuasive effects: A review of current approaches. In S. Jackson (Ed.), *Argumentation and values: Proceedings of the ninth Alta conference on argumentation* (pp. 88–92). Annandale, VA: Speech Communication Association.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Mateo, CA: Morgan Kaufman.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to Persuasion.. *Journal of Personality and Social Psychology, 46*(1), 69–81.

Petty, R. E., & Wegener, D. T. (1991). Thought systems, argument quality, and persuasion. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 4, pp. 147–161). Hillsdale, NJ: Erlbaum.

Pratt, J., Raiffa, H., & Schlaifer, R. (1995). *Introduction to statistical decision theory*. Cambridge, MA: MIT Press.

Putnam, H. (2002). *The collapse of the fact/value dichotomy*. Cambridge, MA: Harvard University Press.

Putnam, H. (2003). For ethics and economics without the dichotomies. *Review of Political Economy, 15*(3), 395–412. doi:10.1080/0953825032000086595.

Railton, P. (2000). Normative Force and Normative Freedom: Hume & Kant. In J. Dancy (Ed.), *Normativity*. Oxford: Blackwell.

Ramsey, F. P. (1931). *The foundations of mathematics and other logical essays*. London: Kegan, Paul, Trench, Trubner & Co.

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.

Robbins, L. (1932). *An essay on the nature and significance of economic science* (2nd ed.). London: Macmillan.

Rosenkrantz, R. D. (1992). The justification of induction. *Philosophy of Science*, 527–539.

Rowbottom, D. P. (2007). The Insufficiency of the Dutch Book argument. *Studia Logica, 87*, 65–71.

Schreier, M., Groeben, N., & Christmann, U. (1995). That's not fair! Argumentational integrity as an ethics of argumentative communication. *Argumentation, 9*, 267–289.

Schroeder, M. (2009). Means-end coherence, stringency, and subjective reasons. *Philosophical Studies, 143*(2), 223–248. doi:10.1007/s11098-008-9200-x.

Schum, D. A. (1994). *Evidential foundations of probabilistic reasoning*. New York: John Wiley.

Satz, D., & Ferejohn, J. A. (1994). Rational choice and social theory. *Journal of Philosophy, 91*(2), 71–87.

Sen, A. (1987). *On ethics and economics*. Malden, MA: Blackwell Publishing.

Sibler, D. S. (1999). Dutch books and agent rationality. *Theory and Decision, 47*, 247–266.

Siegel, H., & Biro, J. (1997). Epistemic normativity, argumentation & fallacies. *Argumentation, 11*, 277–292.

Siegel, H., & Biro, J. (2008). Rationality, reasonableness, and critical rationalism: Problems with the Pragma-dialectical view. *Argumentation, 22*, 191–203.

Simon, H. A. (1982). *Models of bounded rationality* (Vols. 1, 2). Cambridge, MA: MIT Press.

Simon, S., Erduran, S., & Osborne, J. (2002). Enhancing the quality of argumentation in school science. In *Proceedings of the annual meeting of the National Association for Research in Science Teaching*, New Orleans, USA.

Skyrms, B. (1993). A mistake in dynamic coherence arguments?. *Philosophy of Science, 60*, 320–328.

Skyrms, B. (1996). *Evolution of the social contract*. New York, NY: Cambridge University Press.

Skyrms, B. (2009). Evolution and the social contract. In *The Tanner lectures on human values 28*. Salt Lake City: University of Utah Press.

Slob, W. H. (2002). How to distinguish good and bad arguments: Dialogico-rhetorical normativity. *Argumentation, 16*, 179–196.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgement. *Organizational Behavior & Human Processes, 6*, 649–744.

Slovic, P., & Tversky, A. (1974). Who accepts savage's axiom?. *Behavioral Science, 19*(6), 368–373.

Snow, P. (1998). On the correctness and reasonableness of Cox's Theorem for finite domains. *Computational Intelligence, 14*, 452–459.

Snow, P. (2001). The reasonableness of possibility from the perspective of cox. *Computational Intelligence, 17*, 178–192.

Sober, E. (1993). *The nature of selection: Evolutionary theory in philosophical focus*. Chicago: University of Chicago Press.
Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Hillsdale, NJ: Lawrence Erlbaum.
Starmer, C. (2005). Normative notions in descriptive dialogues. *Journal of Economic Methodology, 12*(2), 277–289. doi:10.1080/13501780500086206.
Stich, S. P. (1985). Could man be an irrational animal?. *Synthese, 64*, 115–135.
Stich, S.P. (1990). *The fragmentation of reason*. Cambridge, MA: MIT Press.
Sugden, R. (2001). The evolutionary turn in game theory. *Journal of Economic Methodology, 8*(1), 113–130. doi:10.1080/1350178001002328.
Thaler, R. H., & Mullainathan, S. (2008). *Behavioral economics. The concise encyclopedia of economics* (2nd ed.). Liberty Fund.
Thomson, J. J. (2008). *Normativity*. La Salle, IL: Open Court.
Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky, *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293–215.
van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation—The pragma-dialectical approach*. Cambridge: Cambridge University Press.
van Eeemeren, F. H., Garssen, B., & Meufells, B. (2009). *Fallacies and judgements of reasonableness: Empirical research concerning pragmadialectical discussion rules*. Dordrecht: Springer.
van Eemeren, F. H., Garssen, B., & Meuffels, B. (2012). The disguised abusive ad hominemempirically investigated: Strategic manoeuvring with direct personal attacks. *Thinking & Reasoning, 18*(3), 344–364. doi:10.1080/13546783.2012.678666.
van Enschot-Van Dijk, R., Hustinx, L., & Hoeken, H. (2003). The Concept of Argument Quality in the Elaboration Likelihood Model. In F. H. van Eemeren, J. Anthony Blair, C. A. Willard, & A. Francisca Snoeck Henkemans, *Anyone who has a view: Theoretical contributions to the study of argumentation* (pp. 319–333). Dordrecht: Kluwer Academic Publishers.
Von Neumann, J., & Morgenstern, O. (1947). *The theory of games and economic behaviour* (2nd ed.). Princeton: Princeton University Press.
Waidacher, C. (1997). Hidden assumptions in the Dutch book argument. *Theory and Decision, 43*, 293–312.
Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*, 273–281.
Williams, B. (1985). *Ethics and the limits of philosophy*. London: Fontana Press.
Woods, J., Irvine, A., & Walton, D. (2004). *Critical thinking, logic & the fallacies*. Toronto: Prentice Hall.
Wyer, R. S., Jr., & Goldberg, L. (1970). A probabilistic analysis of the relationships among beliefs and attitudes. *Psychological Review, 77*(2), 100–120.