Formal Epistemology

Igor Douven
Sciences, Normes, Décision, Paris-Sorbonne University
igor.douven@paris-sorbonne.fr

Jonah N. Schupbach
Department of Philosophy, University of Utah
jonah.n.schupbach@utah.edu

Formal epistemology is a flourishing subfield of analytic philosophy characterized both by its matter and method. Its subject matter is epistemology, the theory of knowledge. Its method for investigating this subject matter involves the use of formal, logicomathematical devices. Formal epistemologists attempt to break new ground on traditional epistemological questions using an ever expanding and improving set of such devices. And the philosophical application of various formal devices has itself given rise to a host of new, hotly debated epistemological questions. In this entry, we begin by discussing the discipline of formal epistemology, its historical background, and foundations (Section 1). Then, we summarize some recent work in formal epistemology, both as it pertains to traditional epistemological puzzles (Section 2) and puzzles born out of the development of formal epistemology (Section 3).

1. Historical background and foundations. As an identified and self-contained subfield of philosophy, formal epistemology is a relative newcomer in analytic philosophy. We are not aware of any references to formal epistemology in the literature before 1990. But in a short time, the field has firmly established itself, with many highly active practitioners, philosophy departments advertising jobs explicitly in formal epistemology, and numerous workshops and conferences every year being exclusively devoted to formal epistemology or specific topics in formal epistemology. Like most academic disciplines, philosophy has its fads and fashions. But formal epistemology, we can now say with some confidence, is not among those: formal epistemology, it appears, is here to stay.

While a new branch, formal epistemology grew on a tree with deep roots. First and foremost, this is true because many of the problems it deals with come from traditional epistemology, which dates back to the ancient Greek philosophers. But even its hallmark methodological approach has its origins in the history of philosophical and mathematical thinking—witness Aristotle's development of formal logic and his application of this tool to the characterization of epistemological concepts like scientific knowledge, or *epistêmê*.

Formal epistemology combines tools taken from logic and probability theory. The logic referred to here is the still relatively modern type of logic that was initiated by Frege, Russell, and other mathematicians working at the end of the nineteenth/beginning of the twentieth century. This logic saw important developments during much of the last century, certainly until the 1970s. Probability theory dates back further, with Huygens and the Port Royal logicians (in collaboration with Pascal) laying its groundwork in the mid-seventeenth-century. Bernoulli (Jacob) and Bayes and Laplace made further important contributions in the eighteenth century, and with de Finetti [1937/1964] and Kolmogorov [1950], probability theory received its contemporary form.

Like analytic philosophers generally, epistemologists have always relied on logic for clarifying or checking their arguments. But it was only after a proper semantics had been developed for modal logic (mainly in the work of Kripke) that they started using logic to *analyze* epistemological concepts. Following Hintikka's pioneering work in his *Knowledge and Belief* ([1962]), we have seen the rise of what is sometimes called "modal epistemology," which seeks to analyze knowledge, justification, and related notions in terms of what goes on, not just in the actual world, but also in various non-actual worlds (typically worlds that are, in some sense, "close" to the actual one).

Mostly, however, the formal part of the work stopped with those definitions, and much of what goes under the name of "modal epistemology" is best classified as belonging to mainstream, rather than to formal, epistemology. Some philosophers did go on to develop formal models of knowledge, justification, and belief using various modal logics. Many recognized that these formal models make highly idealizing assumptions about knowledge, or justification, or whichever epistemic notion or notions they intend to represent—such as the "epistemic closure" principle that agents know all of the propositions logically entailed by the propositions they know (which is validated by Hintikka's epistemic logic), or that everything that is believed is also believed to be believed (the positive introspection principle, as validated by KD45 and kindred logics of belief; see Meyer and van der Hoek [1995]). On the other hand, we know from the sciences that models that make idealizing assumptions can still be predictively accurate or valuable in other ways.

Nevertheless, epistemic logics have never gained as much traction in formal epistemology as probabilistic approaches. This is arguably because probability theory offers a modeling tool that is not only more versatile and flexible, but that also does justice to the insight that, to understand how humans cognitively relate to the world, we must attend to the fact that our reasoning and thinking more often than not proceeds on the basis of uncertain premises—premises we believe not categorically but only to a degree—and that it results, more often than not, in our arriving at uncertain conclusions.

The insight that a full understanding of human thinking and rationality requires taking seriously a graded notion of belief received much of its impetus from work in psychology starting in the 1980s. Until then, it had been the received view among psychologists that good reasoning is a matter of obeying the laws of logic. But logic was developed to facilitate mathematical reasoning, in which we go from axioms to theorems via inferential steps that are necessarily truth-preserving. Psychologists have noticed that much of our non-mathematical reasoning can be good, despite being uncertain and defeasible. Accordingly, they claim that the standards of rationality for such reasoning are not provided by a monotonic logic but must be sought elsewhere. In work that marked the beginning of what is now generally known as "the New Paradigm" (Over [2009], Elqayam and Over [2013]), psychologists discovered that people did, overall, quite well at probabilistic reasoning, despite the fact that they were also prone to commit certain fallacies, as had in fact already been reported in earlier work (e.g., Kahneman, Slovic, and Tversky [1982]).

Because of its centrality to formal epistemology, it is worth being clear about what probability theory is. At bottom, the theory is quite simple: all there is to it are a couple of easy-to-grasp axioms plus a definition. These are the axioms:

A1. $0 = \Pr(\varphi \wedge \neg\varphi) \leqslant \Pr(\varphi) \leqslant \Pr(\varphi \vee \neg\varphi) = 1$;
A2. $\Pr(\varphi \vee \psi) = \Pr(\varphi) + \Pr(\psi) - \Pr(\varphi \wedge \psi)$.

In words, these axioms state that all probabilities are between 0 and 1, with logical falsehoods receiving a probability of 0 and logical truths receiving a probability of 1; and that the probability of a disjunction is the sum of the probabilities of the disjuncts minus the probability of their conjunction. The definition concerns the notion of conditional probability, the probability of one proposition *given*, or *on supposition of*, the truth of another proposition. This definition states that the probability of φ given ψ, $\Pr(\varphi \mid \psi)$, equals the probability of the conjunction of φ and ψ divided by the probability of ψ, so $\Pr(\varphi \wedge \psi)/\Pr(\psi)$. Someone whose graded beliefs are representable by a probability function—a function $\Pr(\cdot)$ satisfying A1 and A2—is said to be *statically coherent*.[1] (This notion of probabilistic coherence is unrelated to the notion of coherence that is more commonly used in mainstream epistemology and which will be discussed below.)

Probability theory is silent on how one's graded beliefs ought to change over time as new information comes in. By far the most formal epistemologists (though not all, as will be seen) have committed themselves to a further principle of *dynamic coherence*, according to which we rationally revise our graded beliefs after learning with certainty new information φ precisely if, for every proposition ψ, our new unconditional graded belief in ψ equals the degree to which we believed ψ conditional on φ before we learned about the truth of φ. Formally, where $\Pr(\cdot)$ and $\Pr_\varphi(\cdot)$ denote the probability functions representing our graded beliefs before and after the receipt of φ, respectively, dynamic coherence requires that $\Pr_\varphi(\psi) = \Pr(\psi \mid \varphi)$, for any ψ. For instance, suppose that in the morning you believe to a degree of .5 that it will rain in the evening, and you also

[1] Some philosophers and psychologists hold that static coherence requires more than obedience to the probability axioms. For instance, some believe that it also requires obedience to some principle that connects graded belief to objective probability, like Lewis' [1980] Principal Principle (see Pettigrew [2015] for more recent discussion of the graded belief–objective probability link). Some researchers also hold that rational reasoners must obey a version of Keynes' [1921] Principle of Indifference, according to which one ought to be equally confident in each of a set of mutually exclusive and jointly hypotheses, absent reasons to the contrary. See for recent discussion Decock, Douven, and Sznajder [2016] and Hawthorne et al. [2016] (the latter paper discusses an interesting possible connection between the Principal Principle and the Principle of Indifference).

believe to a degree of .8 that it will rain in the evening on the condition that the afternoon will be cloudy. Then later, when you see that the afternoon is cloudy (assuming you have learned no other relevant information besides this), you must, on pain of being dynamically incoherent, change your graded belief for rain in the evening to .8. If you do, then you are said to update via *Bayes' rule*, also known as the rule of (strict) *conditionalization*.[2]

2. Formal approaches to mainstream and traditional epistemology. In this section, we briefly summarize a few of the ways that formal epistemologists have confronted important traditional and mainstream epistemological questions.

Internalism and Externalism

What does it take for an agent $\mathscr{A}$ to have an epistemically justified belief? Mainstream epistemologists famously divide on this question into two general camps, the internalist and externalist. On the one hand, internalists emphasize the first-person perspective of an epistemic agent. Ask yourself: What does it take for me to be epistemically justified in believing something? The most compelling and common answers assert that I have a justified belief when this belief rests upon my internally having sufficient evidence or reason for that belief. Further questions immediately arise. In what sense must I *have* the requisite evidence or reason? When can my beliefs be said to "rest upon" specific grounds? How much evidence suffices for epistemic justification? etc.

By contrast with the internalist's "egocentric" notion of justification, externalists motivate their "naturalized" accounts by emphasizing a third-person perspective. Additionally, externalists typically place their focus on the concept of knowledge rather than justification. The most salient question for the externalist is: under what conditions would we allow that an epistemic agent has knowledge? The most compelling and common answers assert that knowledge comes by way of their being the right sort of natural relation between the agent's belief state and the world. Justification, when mentioned at all by the externalist, is often taken generically to stand for whatever external relation turns true belief into knowledge. Again, questions immediately arise. Most notably, what precisely constitutes the right sort of natural relation?

Formal epistemologists investigate all of the above questions. For example, some basic work in epistemic logic helps focus the internalism / externalism debate by articulating precise principles bearing on the having of justification and knowledge. In epistemic logic, one adopts a standard Kripkean modal logic but reinterprets the modal operators epistemically. The salient notion of necessity is meant to refer to the state of

[2] Advocates of Bayes' rule acknowledge that not all learning consists of coming to know with certainty the truth of some proposition. Sometimes we just become *more* certain of a proposition, without becoming *entirely* certain of it. Jeffrey [1965] proposed a rule for the accommodation of this type of learning event. The rule—now commonly referred to as "Jeffrey's rule"—states that it ought to hold for all φ and ψ that $\mathrm{Pr}_{\mathrm{new}}(\varphi) = \mathrm{Pr}_{\mathrm{old}}(\varphi \mid \psi) \times \mathrm{Pr}_{\mathrm{new}}(\psi) + \mathrm{Pr}_{\mathrm{old}}(\varphi \mid \neg\psi) \times \mathrm{Pr}_{\mathrm{new}}(\neg\psi)$, with $\mathrm{Pr}_{\mathrm{old}}(\cdot)$ and $\mathrm{Pr}_{\mathrm{new}}(\cdot)$ representing one's graded beliefs before and, respectively, after the learning event that shifts one's confidence in φ. (Note that Bayes' rule falls out of Jeffrey's rule as the special case in which we become certain of ψ.)

knowing—or what *must* be the case, given what we know. The corresponding notion of possibility is meant to refer to the state of *not knowing not*—or what *might* be the case, for all we know. To mark these distinct interpretations, we replace the standard modal operator $\Box$ with the more suggestive K.

One's choice of epistemic logic is directed by reinterpreting and evaluating standard modal axioms. Most notably for present purposes, modal logic's S_4 axiom is restated as $K\phi \rightarrow KK\phi$ and reinterpreted as requiring that one knows some ϕ only if one knows that one knows ϕ. This controversial axiom has come to be known as the KK principle.[3]

The KK principle is sometimes taken as a formally exact criterion dividing internalists and externalists. Anyone who accepts this thesis will require that knowledge has a "luminosity" (Williamson [2000]) about it such that knowers of ϕ not only believe themselves to know ϕ, but they always believe this correctly in whatever way it takes for that true belief to be a case of knowing that they know ϕ. Various forms of internalism naturally motivate the idea that knowledge is so luminous. For example, perhaps knowledge of ϕ requires true belief in ϕ plus some justification-involving "warrant" condition; and perhaps epistemic agents have internal, reflective access to facts about whether their beliefs are warranted. If I (the epistemic agent) know any ϕ in this sense, this internal access naturally (always?) provides me with a warranted new meta-belief that: I am warranted in believing ϕ. Since I believe ϕ itself, I will also then know (i.e., truly believe, with the requisite warrant) that I know ϕ. Even granting the above form of internalism, the argument to the KK principle is far from air tight; more details would be needed to fully motivate the principle. Nonetheless, while internalism does not imply the KK principle, it does seem like the sort of philosophical framework needed to make this principle at all appealing.

On the other hand, it seems clear that externalists will want to reject the KK principle. Whatever the nature of the external relation between belief state and world that must be satisfied to convert true beliefs into knowledge, it is (*qua* externalist) not a relation that agents need be internally aware of satisfying. Indeed, it is tempting to characterize externalism about knowledge in such a way that it straightforwardly entails the denial of the KK principle: externalist accounts of knowledge allow that an agent can know some ϕ without believing (let alone knowing) that they satisfy the externalist condition for having this knowledge—and thus without believing (let alone knowing) that they have this knowledge. Although the question of how the KK principle bears on the internalism / externalism debate is still being explored (e.g., see Okasha [2013]), externalism is arguably much closer to implying the falsity of the KK principle than internalism is to implying its truth—in work in progress, Bird and Pettigrew ([2016]) uncover a precise sense in which this is true.

[3]Relatedly, the S_5 axiom becomes a "principle of negative introspection" $\neg K\phi \rightarrow K\neg K\phi$, requiring that one knows about all the cases in which they fail to have knowledge. Regardless of whether one is an internalist or externalist, this principle is dubious. There are many ways one might fail to have knowledge of some ϕ. Least controversially, ϕ may be believed in all the right ways but false. This principle would require one to know about all such cases.

This is one example of how epistemic logic can be used to clarify precise principles over which internalists and externalists clash. Marking such exact points of disagreement helps to pinpoint what is at stake when internalists and externalists differ over whether justification (or the grounds of knowledge) must be internally had by a knower. In this way, a formal approach focuses the debate over the nature of justification.

In addition to epistemic logic, formal epistemologists have used probability and statistics to investigate issues bearing on internalism and externalism. This is a natural move given that concepts like justification and evidential support are thought of by contemporary epistemologists as defeasible, gradational, and fallible.[4]

Prima facie, an internalist theory of justification is nicely explicated with a Bayesian logic. Probabilities, for the Bayesian, are inherently subjective at least in the sense that they are interpreted as a particular epistemic subject's "degrees of (rational) belief". Bayesians require that an agent's degrees of belief be *statically* and *dynamically coherent* (see Section 1). The epistemic agent has a stock of "background knowledge", and that agent's degrees of belief, to be (statically and dynamically) coherent, must be *fixed* by these known propositions in such a way that they satisfy the axioms of probability A1 and A2. Bayes's rule (along with generalizations thereof) provides the Bayesian with a constraint on how an agent's knowledge (new evidence or reasons) fixes rational degrees of belief.

This all does have an internalist ring to it. An agent's own credences are the locus of attention on this account, and these seem plainly internal to the agent.[5] Any rational change to degrees of belief is affected by the agent "learning" new evidence; formally, via Bayes's rule, we add such evidence to the agent's background knowledge. This change too seems internal to the agent's cognition. But where exactly in the formal framework should the Bayesian locate the notion of internalist justification? The answer seems to depend on what more the Bayesian may or may not want to say about how rational degrees of belief are constrained.

The straightforward answer seems to be that internalist justification is a matter of degree, measured by the Bayesian probabilities themselves; Bayesian degrees of rational belief just are degrees of justification. But this tempting idea at best only makes sense on an "objective Bayesian" account, according to which all rational credences are fixed at precise real values by background knowledge and logical principles.[6] The majority

[4]Observing that these features fit well with a notion of probabilistic support, mainstream epistemologists today often refer to justification as "probabilistic". For example, Fumerton writes, "To be justified in believing one proposition P on the basis of another proposition E, one must be (1) justified in believing that E and (2) justified in believing that E makes probable P" ([1995]:36); Klein also refers to justification as probabilistic when he weakens condition (2), requiring only "that it be *true* that E makes probable P" ([1998]:923).

[5]Even if their *accessibility* may not be as plain. In fact, in what sense (if any) and to what extent degrees of belief are accessible, is a longstanding source of major controversy in formal epistemology (Ramsey [1926/1990], Bradley [2001]).

[6]Even in the complete absence of background knowledge, objective Bayesians hold that there is exactly one unique rational degree of belief that everyone ought to have—on pain of irrationality—in any particular proposition. In such a case, logical principles alone fix rational degrees of belief. It should be noted that the additional logical principles of rationality defended by objective Bayesians often clash with

of Bayesians take a "subjective" stance to some extent, allowing that rational degrees of belief are not entirely fixed in all cases. For example, the most subjective of Bayesians assert that rationality requires no more than that degrees of belief be statically and dynamically coherent. But two hypothetical epistemic agents sharing all of the same background knowledge could differ wildly in all of their probability assignments while both being statically and dynamically coherent. Similarly, two agents could have identical probability assignments in all propositions, but differ greatly in how much background knowledge they have supporting these credences.

An example makes the problem clear. Two epistemic agents are confronted with a coin. Bill is a completely naive agent who has no information about the behavior of this coin, while Hazel has experimented tirelessly flipping this coin. Let us say that Hazel has observed exactly $750,000$ heads and $250,000$ tails in $1,000,000$ flips of the coin. Bill knows that the coin will be flipped again and knows that it has to land heads or tails, but he has no reason whatever to think it will land one way over the other (Bill has not discussed the matter with Hazel, experimented with or even examined the coin, etc.). Nonetheless, Bill and Hazel have equal degrees of belief with respect to this coin landing heads (H) versus tails (T) on the next flip: $Pr(H) = .75, Pr(T) = .25$. Both agents are statically coherent, and we may stipulate that they are both dynamically coherent as well. Shall we allow that they are both equally justified (to degree .75) in believing that the coin will land heads on the next flip? Of course not; Hazel has far more evidence for believing this and so is far better justified.

This motivates another possible answer to the original question for the subjective Bayesian. The degree to which an agent is justified in having some degree of belief is not measured by the degree of belief itself but in the amount of background evidence serving to fix this credence. The subjective Bayesian can (and should) distinguish an agent's credence from the "weight of evidence" grounding that credence. This weight of evidence is what plausibly explicates degree of justification for the subjective Bayesian—though how to formalize weight of evidence is a matter of ongoing investigation (Joyce [2005]).

With a Bayesian explication of justification in hand, internalists may gain a new foothold on resolving some venerable problems. Most famously, given the internalist's notion of justification, we might doubt with Hume ([1748/1912]) whether beliefs about the world ("matters of fact") can ever really be "founded on reasoning, or any process of the understanding." But some formal epistemologists argue that the probability theory provides us with new tools allowing us to respond to Hume's problem of induction (Earman and Salmon [1992], Howson [2000], McGrew [2001]).

Even so, general complications remain for a Bayesian explication of internalist justification. First, note that weight of evidence really at best explicates the degree to which a credence is justified rather than any proposition or propositional belief. If a vast amount of evidence leads an agent's credence to converge on $Pr(P) = .5$, the weight of evidence might be very great indeed. But this does not mean that the agent is strongly justified

the standard *dynamic coherence* constraint. Strictly speaking, objective Bayesians who accordingly reject *dynamic coherence* do not adopt the fundamental Bayesian viewpoint.

in believing P. In fact, the agent seems just as justified in believing $\neg P$, which is just as strongly favored by all of the evidence in hand (since $Pr(\neg P) = 1 - Pr(P) = .5$). The agent is however strongly justified in having this credence. So it is the justification of credences rather than beliefs or propositions that is being explicated by weight of evidence; the explicandum has shifted. Second, the idea that Bayesianism may nonetheless provide a formal internalist epistemology of degrees of belief has not gone unchallenged. Recently, Goldman ([2009]) and Meacham ([2010]) independently (and all too briefly) argue that Bayesianism cannot explicate a *pure* form of internalism, since it requires external constraints on rational degrees of belief.

Can probability and statistics serve to illuminate and develop externalist accounts? This question is receiving increasing attention by formal epistemologists in recent years. To follow one example, Roush ([2005]) uses the probability theory to reexamine and develop Nozick's ([1981]) original counterfactual, tracking theory of knowledge. This theory analyzes the concept of epistemic agent S knowing proposition P as S having true belief in P while satisfying two externalist conditions:

Sensitivity. If P were not true, then S would not believe it.
Adherence. If P were true, then S would believe it.

Roush replaces Sensitivity and Adherence with the following probabilistic conditions:[7]

1. The probability that a subject does not believe P given P is false, $Pr(\neg B(P)|\neg P)$, is greater than threshold t, where $.95 < t < 1$.
2. The probability that a subject believes P given that P, $Pr(B(P)|P)$, is greater than threshold s, where $.95 < s < 1$.

The effect is that Roush's development of the tracking theory sidesteps a host of counterexamples to Nozick's theory. For example (Roush [2005]:98-100), Sam sees Judy and believes P: that the girl he just saw is Judy. However, he is unaware of the fact that Judy's identical twin Trudy was nearby, and that it was only by a twist of luck that he saw Judy herself and not Trudy—in which case Sam would have believed P falsely. In the nearest $\neg P$ worlds, Sam may not see Judy or Trudy, in which case his belief satisfies Sensitivity (and Adherence). But contrary to the original tracking account, few epistemologists would allow that Sam's true belief is knowledge. Roush's account gets this right. Condition 1 is broken, since there is a significant chance that Sam sees Trudy and comes to believe P falsely (after all, it was a mere twist of luck that this didn't happen); $Pr(B(P)|\neg P)$ is not low, and so $Pr(\neg B(P)|\neg P)$ is not sufficiently high.

Several formal epistemologists offer alternative formal developments of the tracking account. Arló-Costa and Parikh ([2006]) reject the approach used by Roush and instead develop a formal tracking account using doxastic logic. Zalabardo ([2012]) offers a probabilistic account in the same basic spirit as Roush's. Kelly ([2014]) proposes a very different computational account of knowledge inspired by tracking accounts.

[7] Roush further develops this account with a recursive definition of knowledge in order to impose closure of knowledge under known implication. Arló-Costa ([2006]) argues that her account still breaks with some intuitive epistemic closure principles, however.

In work in progress, Mayo-Wilson ([2016]) also proposes an alternative tracking account, but then goes further and tests his formal account against scientific practice. As he writes, "Science is widely considered a source of knowledge *par excellence*. So it would be an unwelcome consequence if the most popular philosophical theories of knowledge entail that science fails to produce knowledge." Using his own modal-probabilistic formalization of Sensitivity and Adherence as conditions for having knowledge, Mayo-Wilson demonstrates that classical statistical methods employed across the sciences effectively produce knowledge. This is a positive result, supporting both the statistical methods used by scientists (insofar as we want these to satisfy Mayo-Wilson's Sensitivity and Adherence conditions), and theories of knowledge built upon such formally precise, externalist requirements (insofar as we desire our definitions of knowledge to allow science to be knowledge-producing).

What can formal epistemology offer the mainstream epistemologist with respect to the internalism / externalism debate? While the answer is ultimately yet to be decided, the above discussion allows us already to say the following. Formal approaches serve to clarify this debate by uncovering exact points of agreement and disagreement between the sides. And the internalist and externalist positions themselves are being clarified; exactly what forms of internalism and externalism there are on offer is made clearer by formal epistemologists striving to get precise about what various accounts of justification and knowledge are claiming. Moreover, we have seen that particular problems and counterexamples for mainstream accounts are finding potential resolution in formally subtler developments of these accounts. To the extent that formal methods provide a fruitful means for developing both internalism and externalism, it may even turn out that formal epistemology doesn't so much advance this longstanding debate as show that there was something to each side's position all along. Justification, after all, is plausibly polymorphous, and there may well be interesting philosophical aspects of internalist and externalist notions of justification (Staley and Cobb [2011]). All in all, in a variety of ways, formal epistemology seems to be offering promising new approaches to investigating these mainstream issues.

The structure of justification

Granting that we can have epistemically justified beliefs, what structure does this justification take? Epistemologists often clarify this question through a famous and ancient puzzle known as the Regress Problem. Say that belief b is epistemically justified because it follows in the relevant way from b_1. This reason must itself be justified if it is to play the requisite role, and so we may press for the justification of b_1. If b_2 is the reason which suffices for justifying b_1, then we may continue our inquiry and find that it is b_3 that justifies b_2, b_4 which justifies b_3 and so on. A regress of reasons naturally arises, and now the question is how this story should end. Three alternate endings suggest themselves:

Infinitism. The regress never stops; an infinite chain of ever new reasons describes the structure of justification.

Coherentism. The regress is really part of a larger network of logical interconnections between propositions; it is the coherence of this larger structure that bestows justification on any of its members.

Foundationalism. The regress eventually terminates in reasons that are justified, but not on the basis of other reasons; all beliefs are ultimately justified by being grounded in these "properly basic" beliefs.

The Regress Problem has often been cited in support of foundationalism (McGrew and McGrew [2008]). A simple version of such an argument takes the form of a reductio. Since legitimate epistemic justification couldn't possibly come by way of infinite chains of reasons, or entirely by way of (possibly circular) interconnections between reasons, rational regresses must eventually terminate in a foundational layer of properly basic beliefs. But formal epistemologists have breathed new life into the infinitist and coherentist alternatives by exploring more carefully whether infinitist and coherentist notions of justification are really so absurd.

In a series of recent publications, Peijnenburg ([2007], [2010]) and Peijnenburg and Atkinson ([2008]) offer a probabilistic investigation into infinitism. Construing justification as a matter of degree explicated by probabilistic support, they ask the following question [2008:333]: "Consider [a] chain $S_0, S_1, S_2, \ldots$, etc., where each S_{n+1} probabilistically justifies S_n. Are we able to justify S_0, in the sense that we can compute $Pr(S_0)$?" The thought is that, since the sequence $\langle S_0, S_1, S_2, \ldots \rangle$ is countably infinite, infinitism's opponents would either want to say that $Pr(S_0)$ must be zero or just undefined. To the contrary, Peijnenburg and Atkinson demonstrate that, under certain conditions, this probability can be calculated.

The demonstration in Peijnenburg [2007] runs as follows: For simplicity, allow that $Pr(S_n|S_{n+1}) = \alpha$ and $Pr(S_n|\neg S_{n+1}) = \beta$ for all n (Peijnenburg later shows that this assumption is not crucial to the demonstration). Additionally, given that S_{n+1} is the probabilistic justifier of S_n (for each n), assume that $\alpha > \beta$. Then, by the rule of total probability,

$$Pr(S_n) = Pr(S_{n+1})\alpha + Pr(\neg S_{n+1})\beta = Pr(S_{n+1})(\alpha - \beta) + \beta$$

Thus, for example,

$$Pr(S_1) = Pr(S_2)(\alpha - \beta) + \beta.$$

And so,

$$Pr(S_0) = Pr(S_1)(\alpha - \beta) + \beta = [Pr(S_2)(\alpha - \beta) + \beta](\alpha - \beta) + \beta$$

Generalizing this calculation to any n, Peijnenburg derives the following equation:

$$Pr(S_0) = \frac{\beta}{1 - \alpha + \beta} + (\alpha - \beta)^{n+1}[Pr(S_{n+1}) - \frac{\beta}{1 - \alpha + \beta}] \tag{1}$$

As Peijnenburg points out, as n increases without bound, the term $(\alpha - \beta)^{n+1}$ occurring in equation (1) becomes smaller and smaller. The result is that, in the limit with n going to infinity, we need not know the value of $Pr(S_{n+1})$ in order to derive $Pr(S_0)$, since the latter is only a function of known values α and β:

$$Pr(S_0) = \frac{\beta}{1 - \alpha + \beta}$$

Here is a case then in which we finite beings *can* calculate $Pr(S_0)$, and in which $Pr(S_0)$ need not be 0, despite S_0 resting on an infinite sequence of reasons. Peijnenburg and Atkinson take themselves to have shown that, under certain conditions, it is possible "to justify S_0, in the sense that we can compute $Pr(S_0)$" ([2008]:337).

What can we say about the significance of this result for the structure of justification debate. Gwiazda [2010] takes issue with the fact that Peijnenburg and Atkinson's result has to do with computability and not "completability"; he asserts that it is the latter which infinitist justification would require, but "there is a wide gulf between demonstrating computability [...] and demonstrating completability." As Peijnenburg ([2010]) convincingly argues in response, however, it is not at all clear that infinitism requires anything beyond computability—nor is the computability / completability distinction clear to begin with.

Even allowing that it speaks appropriately to the notion of infinitist justification, one should be careful not to read too much into this formal result. The foremost defender of infinitism today, Peter Klein, clarifies "the essential claim of infinitism" as this: "The reasons that justify a belief are members of a chain (perhaps branching) that is infinitely long and non-repeating" ([1998]:919). Infinitism makes a claim on the general structure of justification, and Peijnenburg and Atkinson's formal work manifestly does *not* establish this general theory (nor do they argue that it does). Still, the result supports infinitism by dismantling the strong intuition shared by many that an infinite string of reasons cannot ultimately provide any reason at all; it does much to show that the infinitist option is not so absurd as many have assumed. It thus takes away any quick argument from the impossibility of infinitist justification to foundationalism (or coherentism).

The structure of justification proposed by the infinitist is straightforwardly explicated, even if it remains doubtful that infinitism describes the general structure of justification. However, when one turns to consider coherentism, it is not even clear what the positive proposal is. Epistemologists speak of coherence vaguely as the property beliefs have to the extent that they "hang together", "agree with each other", "dovetail", or "support one another". However, coherentists and their opponents alike bemoan this theory's lack of substance given that there is no clearer explication of coherence; e.g., Bonjour [1985:94] famously writes, "[T]he main work of giving [a general account] which will provide some relatively clear basis for comparative assessments of coherence, has scarcely been begun, despite the long history of the concept." Prior to investigating whether coherentism describes the structure of justification, one must first get a clear grasp on what coherence amounts to.

The most popular formal approach to this problem is again probabilistic, and specifically Bayesian.[8] Shogenji [1999] initiated a flurry of work developing and evaluating Bayesian measures of coherence (Olsson [2002], Bovens and Hartmann [2003], Fitelson [2003], Glass [2005], Douven and Meijs [2007a], Schupbach [2011a], etc.). Shogenji's original proposal is that the coherence of an agent's set of believed propositions $S = \{B_1, B_2, ...B_n\}$ be explicated by comparing the joint probability of the members of S $Pr(B_1 \wedge B_2 \wedge ... \wedge B_n)$ with the value that this joint probability would take were these members statistically independent of one another $Pr(B_1) \times Pr(B_2) \times ... \times Pr(B_n)$:

$$Coh_S(S) = \frac{Pr(B_1 \wedge B_2 \wedge ... \wedge B_n)}{Pr(B_1) \times Pr(B_2) \times ... \times Pr(B_n)}$$

The function *Coh* thereby provides a measure of the degree to which the members of this information set are statistically dependent on, or relevant to, one another. Such a measure thus intuitively captures the idea that coherence is a matter of how well beliefs "dovetail" or "support one another". For simplicity, consider the case where we are evaluating the coherence of a pair of beliefs, $S = \{B_1, B_2\}$. Shogenji's measure can be rewritten as follows:

$$Coh_S(S) = \frac{Pr(B_1 \wedge B_2)}{Pr(B_1) \times Pr(B_2)} = \frac{Pr(B_1) \times Pr(B_2|B_1)}{Pr(B_1) \times Pr(B_2)} = \frac{Pr(B_2|B_1)}{Pr(B_2)} = \frac{Pr(B_1|B_2)}{Pr(B_1)}.$$

These last two ratios straightforwardly measure the degree to which either of the propositions supports the other (as the extent to which the truth of either proposition would make the other more likely).

Despite its intuitive appeal and simplicity, formal epistemologists have put forward various arguments against Shogenji's account and developed alternative measures for evaluation. As one example of how this dialectic goes, Fitelson ([2003]) argues that Shogenji's general measure focuses exclusively on the coherence of all the propositions in set S at once and thus ignores support relations between various of S's proper subsets. Schupbach ([2011a]) develops an example showing that this "Depth problem" does indeed lead Shogenji's explication into counterintuitive results. Both Douven and Meijs ([2007a]) and Schupbach offer "subset-sensitive" developments of Shogenji's original measure that do not fall prey to Depth problem, but other criticisms of these developments have now been published (e.g., Siebel and Schippers [2015]).

[8] For a different formal approach, the reader is directed to Thagard's computational account of coherence as "constraint satisfaction". In any "coherence problem," we hunt for a subset of a set of elements $\mathbf{E} = \{e_1, e_2, ..., e_n\}$ which best satisfies a given set of constraints. These constraints may link elements in a positive way (e.g., if e_j implies e_k, a positive constraint may require that we accept both), or in a negative way (e.g., if e_j is inconsistent with e_k, a negative constraint may require that we accept only one). Each particular constraint is assigned a weight. Thagard ([2000]) explicates coherence as W, the sum of the weights corresponding to the constraints that get satisfied by a particular partition of $\mathbf{E}$; thus, maximizing coherence amounts to partitioning $\mathbf{E}$ into two sets (accepted and rejected) in a way that maximizes W. Thagard ([1989]) develops a program ("ECHO") that computes approximate solutions to explanatory coherence problems using a connectionist (neural network) algorithm.

An interesting example of an early alternative explication of coherence is put forward by both Olsson ([2002]) and Glass ([2005]):

$$Coh_{OG}(S) = \frac{Pr(B_1 \wedge B_2 \wedge ... \wedge B_n)}{Pr(B_1 \vee B_2 \vee ... \vee B_n)}$$

While Shogenji's measure (and measures inspired by it) nicely captures informal descriptions of coherence along the lines of "mutual support", the Olsson-Glass measure (and measures inspired by it) rather target the intuitive idea of coherence as "agreement". Note that Coh_{OG} takes its maximal value 1 when $Pr(B_1 \wedge B_2 \wedge ... \wedge B_n) = Pr(B_1 \vee B_2 \vee ... \vee B_n)$. This occurs just when the various B_i all agree maximally with one another, being logically equivalent. Similarly, Coh_{OG} is minimal when the B_i stand contrary to one another.

The debate over which, if any, proposed measure best explicates the epistemological notion of coherence continues today, but the examples above suggest the possibility that we need not automatically feel compelled to choose between alternatives. The concept of coherence finds precise explication through the work of formal epistemologists. But the formal approach also has the potential to disentangle distinct concepts otherwise conflated under the heading "coherence".

What do such accounts say to the original question of epistemological interest, whether coherentist structures can adequately describe the structure of justification? Of course, ultimately the answer will hinge on which of the precise formal characterizations of coherence one accepts. As with infinitists, coherentists first rephrase the question more precisely: are more coherent sets of beliefs more probably true? Phrased in this way, the answer is clearly negative. In fact, using *any* proposed, Bayesian measure of coherence, one can show that a more coherent of a set of beliefs may even be less probable. Using Shogenji's account, the set of beliefs $S = \{B_1, B_2, ..., B_n\}$ may be more coherent than another set $S' = \{B'_1, B'_2, ..., B'_n\}$:

$$\frac{Pr(B_1 \wedge B_2 \wedge ... \wedge B_n)}{Pr(B_1) \times Pr(B_2) \times ... \times Pr(B_n)} > \frac{Pr(B'_1 \wedge B'_2 \wedge ... \wedge B'_n)}{Pr(B'_1) \times Pr(B'_2) \times ... \times Pr(B'_n)}$$

and simultaneously $Pr(B_1 \wedge B_2 \wedge ... \wedge B_n) < Pr(B'_1 \wedge B'_2 \wedge ... \wedge B'_n)$. These formal accounts go beyond this negative conclusion, however. They also illuminate how and why this can be. In this case, the more coherent belief set can be less probable if $Pr(B_1) \times Pr(B_2) \times ... \times Pr(B_n) \ll Pr(B'_1) \times Pr(B'_2) \times ... \times Pr(B'_n)$; in words, the fact that $S = \{B_1, B_2, ..., B_n\}$ is more coherent than $S' = \{B'_1, B'_2, ..., B'_n\}$ may be outweighed by the fact that the individual beliefs making up S are on average much less plausible than those making up S' (see Klein and Warfield [1994]).

Formal epistemology, in its current state, thus casts doubt on coherence theories of justification. This is, of course, not to say that coherence cannot be thought of as justification-conducive (or an epistemic virtue) in some weaker sense. In fact, another controversy in the contemporary literature on Bayesian Coherentism concerns whether more coherent bodies of belief are demonstrably more probably true, *ceteris* *paribus*—where this "all else being equal" clause may arguably require formal concepts like $Pr(B_1) \times Pr(B_2) \times ... \times Pr(B_n)$ (what Shogenji calls "total individual strength") to be held constant between appropriately comparable sets of beliefs (Bovens and Hartmann [2003], Douven and Meijs [2007b], Schupbach [2008], Huemer [2011]). Nonetheless, work in formal epistemology puts into doubt the thesis that coherence alone can describe the structure of justification generally—at least insofar as justification is thought of probabilistically. It seems rather that coherence is at best one epistemically relevant factor among potentially many that determines the extent to which some set of beliefs is justified.

Finally, the probability theory has also been used to shed new light on various issues for foundationalism. We have already noted that formal work on infinitism might be taken to challenge the foundationalist idea that a terminal layer of properly basic beliefs is necessary if we are to have epistemic justification at all. A similar argument can be made against a "pure" form of foundationalism from the apparent need for circles in the architecture of justification. Haack ([1993]:86; see also Dancy [2003]) puts forward such an argument, asserting that a pure foundationalist theory cannot account for reasoning so mundane as that used for crossword puzzles. Such reasoning is characterized by mutually supporting "circles" of justification: "How reasonable one's confidence is that 4 across is correct depends, *inter alia*, on one's confidence that 2 down is correct [... H]ow reasonable one's confidence is that 2 down is correct in turn depends, *inter alia*, on how reasonable one's confidence is that 4 across is correct."

In response to such arguments, McGrew and McGrew ([2008]:57) take on the foundationalist's onus of "model[ing] the phenomenon of mutual support without violating the principle that circular reasoning is non-justificatory." They proceed by constructing a foundationalist model of scenarios in which two beliefs H_1 and H_2 support one another by filling in underlying foundational justifications for each of these separate beliefs. For their example, they stipulate that H_1 is supported by foundational beliefs F_1 and F_A, while H_2 is justified by foundational beliefs F_2 and F_B. With justification and support construed probabilistically, the mutual support between H_1 and H_2 amounts to noting that $Pr(H_1|H_2) > Pr(H_1)$ iff $Pr(H_2|H_1) > Pr(H_2)$.

At this stage, the appearance of circular support remains. After all, F_1 and F_A jointly support H_1, which supports H_2, which supports H_1 again! Correspondingly, F_2 and F_B jointly support H_2, which supports H_1, which supports H_2 again! To dispel this appearance, McGrew and McGrew must show that the support which either belief receives from the other is somehow separable from the support it bestows on the other; e.g., the support H_1 receives via H_2 is distinguishable from the support H_1 bestows upon H_2. They do so using two formal maneuvers. First, either belief "screens off" its foundational justifiers from the other belief. Put formally:

$$\begin{aligned}
Pr(H_2|H_1 \wedge F_1 \wedge F_A) &= Pr(H_2|H_1) \\
Pr(H_2|\neg H_1 \wedge F_1 \wedge F_A) &= Pr(H_2|\neg H_1) \\
Pr(H_1|H_2 \wedge F_2 \wedge F_B) &= Pr(H_1|H_2) \\
Pr(H_1|\neg H_2 \wedge F_2 \wedge F_B) &= Pr(H_1|\neg H_2)
\end{aligned}$$

Informally, this means that all of the support that H_1 receives from F_2 and F_B comes by way of H_2; in McGrew and McGrew's terminology, "H_2 is a conduit through which F_2 and F_B support H_1" (similarly, H_1 serves as a conduit of support from F_1 and F_A to H_2).

The second formal maneuver allows McGrew and McGrew to model the idea that the *only* support either belief (H_1 or H_2) provides for the other is in its role as such a conduit. They formalize this support using Jeffrey conditionalization (see footnote 2) on the intermediate, conduit beliefs. Continuing with the above example, imagine that F_A and F_B are already background knowledge for an agent, and then the agent comes to learn F_1 and subsequently F_2. Upon learning F_1, the agent updates the probability of H_1—using Bayesian conditionalization—to get $Pr_{\text{new}}(H_1) = Pr_{\text{old}}(H_1|F_1)$. The probability of H_2 is then updated using Jeffrey conditionalization:

$$Pr_{\text{new}}(H_2) = Pr_{\text{old}}(H_2|H_1) \times Pr_{\text{new}}(H_1) + Pr_{\text{old}}(H_2|\neg H_1) \times Pr_{\text{new}}(\neg H_1)$$

Given that the screening off relation holds, this equation can be reduced as follows:

$$\begin{aligned}
Pr_{\text{new}}(H_2) &= Pr_{\text{old}}(H_2|H_1) \times Pr_{\text{new}}(H_1) + Pr_{\text{old}}(H_2|\neg H_1) \times Pr_{\text{new}}(\neg H_1) \\
&= Pr_{\text{old}}(H_2|H_1) \times Pr_{\text{old}}(H_1|F_1) + Pr_{\text{old}}(H_2|\neg H_1) \times Pr_{\text{old}}(\neg H_1|F_1) \\
&= Pr_{\text{old}}(H_2|H_1 \wedge F_1) \times Pr_{\text{old}}(H_1|F_1) + Pr_{\text{old}}(H_2|\neg H_1 \wedge F_1) \times Pr_{\text{old}}(\neg H_1|F_1) \\
&= Pr_{\text{old}}(H_2|F_1)
\end{aligned}$$

The upshot is that the support that H_2 receives from H_1 just amounts to that which it receives from gaining foundational belief F_1; more generally in the above example, similar demonstrations show that the support H_1 provides for H_2 entirely comes by way of foundational beliefs F_1 and F_A, while the support that H_2 provides for H_1 is grounded entirely in F_2 and F_B. On this framework then, the appearance of circular support is illusory, all justification ultimately deriving from foundational beliefs. Note that this work is not meant to be an argument for this foundationalist account, so much as a way of showing that the "pure foundationalist" can make sense of mutual support—without allowing circular reasoning to be justificatory.

Social epistemology

Traditionally, epistemologists have written as though individuals seek knowledge entirely on their own, having to learn about the world in complete isolation of others. Mainstream epistemologists have only recently focused on the question of what difference it might make that, in reality, we socially interact with others who largely have the same epistemic goals as we. In ground-breaking work, Goldman [1999] showed that the exclusive focus on the isolated epistemic agent was deeply mistaken. There are important aspects of our epistemic lives that can only be understood by considering our interactions with our fellow epistemic agents and by studying whole collectives of agents pursuing truth in a concerted effort. By pretending that our being members of such collectives is irrelevant to the study of epistemology we are missing important pathways to justified belief and knowledge: much of what we justifiably believe or know is due to our interacting with others.

This insight led to increased attention investigating the epistemic significance of such topics as testimony, expertise, judgment aggregation, and disagreement. So far, much of the work undertaken on these and related topics is still non-formal and best classified as mainstream epistemology. But here we can briefly mention recent work suggesting the emergence of a subfield within formal epistemology, to wit, formal *social* epistemology.

We sometimes ascribe knowledge or belief to groups. We might say that the Obama administration believes climate change to call for more drastic measures, that the present government knows that researchers are underpaid, or that the court found the defendant guilty. And we may make such claims even if not *every* member of the Obama administration agrees that a more drastic response to climate change is necessary, if some members of the government actually disagree that researchers are underpaid, or if one of the judges dissented with the verdict. So, consensus that φ is seemingly not required to ascribe knowledge of or belief in φ to a collective. On the other hand, we presumably also would not want to ascribe knowledge or belief to the collective if not at least a majority of its members knew or believed that φ. Interestingly, Kornhauser and Sager [1986] have shown that aggregating what individuals know or believe on the basis of a simple majority rule—the collective knows/believes any proposition that most of its members know/believe—can give rise to inconsistencies at the aggregated level even if no individual is inconsistent in any way.

The following is a classical demonstration of the problem: Suppose a committee of three has to decide about whether a certain applicant will be made a job offer, and define φ ≔ "The applicant has a strong publication record"; ψ ≔ "The applicant has sufficient teaching experience"; χ ≔ "The applicant should receive an offer." All three committee members agree that the applicant should receive an offer iff she has a strong publication record and she has sufficient teaching experience, but they do not quite agree on the atomic propositions. Specifically, member 1 believes φ, ψ, and χ; member 2 believes φ but believes both ψ and χ to be false; and member 3 believes ψ but believes φ and χ to be false. One readily verifies that all three members hold consistent beliefs. However, if group belief goes by majority, then the group as a whole is inconsistent. After all, a majority of its members (members 1 and 2) believe φ; a majority (members 1 and 3) believe ψ; *all* believe the biconditional $(\varphi \wedge \psi) \leftrightarrow \chi$; but while this biconditional in conjunction with φ and ψ entails χ, a majority (members 2 and 3) believe that proposition to be *false*.

This discovery motivates various questions: Ought aggregation procedures preserve consistency? Which candidate procedures preserve consistency (and under what conditions)? What other desiderata might we have for a satisfactory method of aggregation? Formal epistemologists debate these questions and more in a growing body of research studying the aggregation of epistemic attitudes. This research has, for example, provided a number of formal possibility and impossibility results, uncovering which aggregation procedures guarantee consistent attitudes (belief, knowledge, but also other propositional attitudes) at the group level under which precise conditions—see List and Puppe [2009], Dietrich and List [2010], and references given there.

The central question of these debates is how we should construct group attitudes on the basis of typically diverging—or at any rate partly diverging—individual attitudes. Whether such divergences in individual attitudes are themselves epistemically significant is the focus of a separate debate in social epistemology, the so-called peer disagreement debate. In particular, this debate has revolved around the question of whether the discovery that one holds a different view on a given matter than one or more of one's peers—people with roughly equal intellectual capacities and access to basically the same evidence—should give one reason to revise that view. According to some participants to this debate, it is perfectly okay to stick to one's own view in the face of peer disagreement, while others hold that such a case calls for revision of one's view, with many believing that the revision should consist of a kind of compromise between one's view and that of one's disagreeing peer or peers. Much of the literature on peer disagreement does not belong to formal epistemology. But both the question of whether peer disagreement is a reason to compromise and the question of how to compromise (supposing that is found to be the right response to peer disagreement) have drawn much attention from formal epistemologists. In fact, proposed compromising models have drawn much inspiration from, and even heavily used, the earlier-cited literature on judgment aggregation. For representative formal work on peer disagreement, see for instance Fitelson and Jehle [2009], Lam [2011], Brössel and Eder [2014], Cevolani [2014], and Levinstein [2015].

If we want to study formal models of compromising in communities of agents, it will generally be difficult to obtain interesting analytic results even if the communities are only moderately large. To study epistemic interactions in such communities, we have been greatly helped by the development of computational environments that allow us to simulate large numbers of interacting agents while keeping track of the relevant epistemic features of those agents. A simulation model that has proven to be particularly helpful in this respect is one developed in Hegselmann and Krause [2002], [2006]. The basic model is very simple and is populated by simulated agents in pursuit of some truth who change their views on what the truth is both on the basis of evidence they receive and on the basis of exchanges they have with some of their fellow agents. Already this simple model was shown to yield interesting results about the conditions under which the views of the agents in the community converge, the conditions under which these views diverge, and much more. Better still, the model is very flexible, making it easy to tweak and extend its basic machinery, and models building on the original model developed by Hegselmann and Krause have been used to study normative questions pertinent to social epistemology (see, e.g., Olsson [2008] and Douven [2010]).

A model for studying epistemically interacting agents that is similar to, but not strictly an extension of, the model developed by Hegselmann and Krause is presented in Olsson [2011], Olsson and Vallinder [2013], and Vallinder and Olsson [2014]. The first paper uses the model for studying, and vindicating, certain theses implied by Goldmann's [1986] reliabilist epistemology. The second paper compares in computer simulations some prominent norms of assertion. And the third paper presents probabilistic models of trust among epistemic agents and shows that, under plausible assumptions, we are often better off putting greater trust in the reliability of our own inquiries than in those of others. For related work, see Sprenger, Martini, and Hartmann [2009] and Hartmann, Pigozzi, and Sprenger [2010].

3. New questions born out of formal epistemology. In this section, we discuss a couple of the epistemological issues that have arisen out of the development of formal epistemology itself.

Connecting categorical and graded belief

With so much of formal epistemology focused on the notion of degrees of belief, questions arise regarding the status of the traditional notion of belief *simpliciter*. With a formally precise, probabilistic concept of credence in hand, some philosophers maintain that any talk of categorical belief is to be shunned as unscientific (most famously Jeffrey [2004]). People have degrees of belief, and probability is crucial to defining what it is for such degrees of belief to be rational and to change in rational ways. But whenever we talk unqualifiedly about what we or others believe, we are talking loosely, and loose talk should not be the subject of serious philosophy. Needless to say, these philosophers take traditional epistemology to be a deeply misguided enterprise.

The majority of formal epistemologists, however, take both notions of belief (categorical and graded) more seriously. There is an epistemology of belief and an epistemology of degrees of belief—as Foley [1992] puts it—and in the view of these formal epistemologists neither is to be dismissed as unscientific or second-rate. Accordingly, many formal epistemologists today wrestle with the question of how degrees of belief and categorical beliefs bear on one another.

Surely graded and categorical beliefs do not just co-exist in our heads, with graded beliefs being completely unconstrained by what we believe categorically, and vice versa. It is thus reasonable to suppose that there must be some connection—some bridge principle or principles, if one likes—between the epistemology of belief and the epistemology of degrees of belief. One answer has been that the rationality of our categorical beliefs supervenes on our rationally held graded beliefs, in the sense that there cannot be a change in the former without there being some change in the latter, and where it is assumed that rationally held graded beliefs are degrees of belief that are representable by a probability function. Indeed, some have suggested the following straightforward connection between graded and categorical beliefs, which is sometimes called "the Lockean Thesis":

Lockean Thesis (LT) It is rational to believe φ (categorically) if and only if it is rational to believe φ to a degree above a certain threshold value θ,

where θ is then typically assumed to be close, but unequal, to 1.

However, (LT) is known to lead to trouble when combined with two principles concerning categorical belief that, at least prima facie, appear difficult to deny, namely, the Conjunction Principle,

Conjunction Principle (CP) It is rational to believe the conjunction of any two propositions that are individually rational to believe,

and the No Contradictions Principle,

No Contradictions Principle (NCP) It is never rational to believe an explicit contradiction.

Specifically, these principles are known to give rise to the so-called lottery and preface paradoxes.

The lottery paradox was first presented in Kyburg [1961].[9] Consider a fair n-ticket lottery $\mathscr{L}$ with exactly one winner, and with $1-1/n > \theta$. Suppose you have been informed about the conditions of $\mathscr{L}$ and consequently believe rationally to a degree exceeding θ that ticket number i in $\mathscr{L}$ is a loser, for all $i\colon\ 1 \leqslant i \leqslant n$. By (LT), it is rational for you to believe (categorically) that ticket number i is a loser, for all i. At the same time, you know, and hence rationally believe, that one of the tickets will be the winner. Now, the conjunction of all those propositions—that ticket number 1 will lose, that ticket number 2 will lose, ..., that ticket number n will lose, *and* the proposition that one of tickets numbers 1 through n will be the winner, forms a contradiction relative to your background knowledge. As a result, you can rationally believe this conjunction, by (CP), but then again you can *not*, by (NCP).

Makinson [1965] was the first to present the preface paradox. This paradox imagines that you have just finished writing a book. You have checked over and over again each of the n claims the book makes, and so you can rationally believe each of those claims to a degree above θ. Thus, you can rationally believe them (categorically, that is), by (LT). However, you are also aware that colleagues spotted errors in your previous books after their publication, despite the fact that you had checked the claims in those books as thoroughly as you checked the claims in your new book, and despite the fact that you were as confident about the correctness of each of the claims in your previous books as you are about the correctness of the claims in the new book. On the basis of this evidence, it would seem rational for you to believe to a degree above θ that the new book will not be entirely free from errors either. Hence, again by (LT), it would seem rational for you to believe categorically that at least one of the claims in the new book is incorrect. But then it is rational for you to believe the conjunction of all the claims contained in the book *and* the proposition that at least one of them is false—which is an explicit contradiction, and according to (NC) you can *not* rationally believe that.

It is a matter of ongoing controversy which of (LT), (CP), and (NCP) is to be abandoned in view of these paradoxes. Here, we can only provide some pointers to the relevant literature: Kyburg [1961], Klein [1985], Foley [1992], Christensen [2004], and

[9] It seems to have been independently discovered by Hempel [1962].

Kroedel [2012] all favor rejecting (CP). Others, including Pollock [1990], Maher [1993], Nelkin [2000], Douven [2002], Wenmackers [2013], and Kelp [2016] believe that (LT) has to go, or at least needs qualification.[10] Priest [1998] has criticized (NCP) on general grounds, and some recent proposals aim to salvage (CP), (NCP), and at least the gist of (LT) by making rational belief a contextual matter, so that it depends on which other propositions one considers whether it is rational to believe of a given ticket that it will lose; see Lin and Kelly [2012], Leitgeb [2014], and Easwaran [Forthcoming].[11]

Formal approaches to explanatory reasoning

There are typically many possibilities (possible worlds, if you like) compatible with our knowledge, some of which appear more likely than others in light of that knowledge. What we do when we receive new information about the world is redistribute our credences across the possibilities. Bayes' rule, as described in Section 1, offers a systematic procedure for such redistributions. As also mentioned above, most formal epistemologists subscribe to Bayes' rule. But in principle there are indefinitely many credence-redistribution procedures.

With the development of such mathematical models comes a set of interesting new questions pertaining to the proper place of less technical theories of inference and rationality. For example, consider the so-called Inference to the Best Explanation (IBE), which in the not-so-distant past enjoyed widespread popularity both among mainstream epistemologists (e.g., Harman [1965] and Vogel [1990]) and among philosophers of science (Boyd [1984], Lipton [2004]). In its crudest formulation, IBE states that we ought to infer the hypothesis that best explains the available evidence. The reference to unqualified inference suggests that IBE belongs to the epistemology of categorical belief. Should IBE be dropped in favor of a formal credence-redistribution rule? Can IBE and various credence-redistribution rules be part of a larger, consistent account of epistemic rationality? Or perhaps can credence-redistribution procedures be developed (or reinterpreted) to validate IBE's central insight that explanatory judgments carry legitimate normative weight? A few philosophers have argued that IBE should effectively be dropped in favor of Bayes' rule or some other allegedly more fundamental model (Fumerton [1980], Salmon [2001]). However, an increasing number of formal epistemologists today are seeking more irenic accounts of how credence-redistribution principles and IBE can share in a full account of rational belief change.

"Bayesian explanationist" approaches attempt to show that Bayes' rule and IBE are somehow compatible (or even mutually supportive) pieces of a general theory of rational inference and belief change. In one version of this approach, formal epistemologists

[10] Especially with regard to proposed qualifications, it is to be emphasized that the lottery and preface paradoxes have nothing specifically to do with lotteries or prefaces, so that limiting applications of (LT) to propositions *not* about lotteries, or propositions that do not assert anything about the correctness of claims made in a book, is not going to be of any help; see Douven and Uffink [2003], Douven and Williamson [2006], Chandler [2010], and Smith [2010].

[11] Chandler [2013] and Briggs et al. [2014] point out various interesting parallels between the lottery paradox and the earlier-cited literature on judgment aggregation.

argue that explanatory judgments are to be used as heuristics for assigning values necessary to running the Bayesian machinery (e.g., Lipton [2004, Ch. 7], McGrew [2003], Huemer [2009], Weisberg [2009], Poston [2014, Ch. 7]). This proposal is not unproblematic. For one thing, the explanationist is not likely to be attracted to a theory that reduces explanatory reasoning's normative import merely to that which it gleans parasitically from the Bayesian framework (see Douven [2011, Sect. 4] and Henderson [2014] for more extensive criticism of the heuristic view).

There are other options, contrasting with the heuristic picture, for how one might develop the general Bayesian explanationist view. For example, one might rather think of IBE and Bayes's rule as situated at different levels of idealization in a unified hierarchy of logics. The idea here would be that by following either Bayes' rule or IBE, epistemic agents (in certain contexts perhaps) tend to reason in accord with the other. For example, by following the less idealized, explanationist logic of IBE, reasoners tend to reason in accord with the more idealized Bayesian logic—under certain conditions, they may reason slightly less reliably and in others more reliably (Schupbach [2016]). Henderson's "emergent compatibilism" ([2014]) and the account defended by Schupbach ([2011b], [2016]) are both along these lines.

Another recent approach is to develop candidate credence-redistribution procedures distinct from Bayes' rule, which give special attention to explanatory relationships between the various possibilities under consideration and the newly obtained information. Here, for example, is a "probabilistic version of IBE" studied in Douven [2013] and Douven and Wenmackers [2016]: Let $\{\psi_i\}_{i\leqslant n}$ be a set of self-consistent, mutually exclusive, and jointly exhaustive hypotheses, and let $\Pr(\cdot)$ represent one's current graded beliefs. Then one probabilistically infers to the best explanation upon learning φ (and nothing else) iff, for all i,

$$\Pr^*(\psi_i) \quad = \quad \frac{\Pr(\psi_i)\Pr(\varphi \mid \psi) + \mathscr{E}(\psi_i, \varphi)}{\sum_{j=1}^{n}\big(\Pr(\psi_j)\Pr(\varphi \mid \psi_j) + \mathscr{E}(\psi_j, \varphi)\big)},$$

with $\mathscr{E}$ assigning a bonus to the hypothesis that explains the evidence best, and with $\Pr^*(\cdot)$ one's new graded belief function. One verifies that this probabilistic version of IBE agrees with Bayes's Rule iff $\mathscr{E}$ is set to be the constant function 0, meaning that no bonus points for explanatory bestness are ever attributed.

One reason to be interested in this and similar alternatives to Bayes' rule is that psychologists have shown that, while people in some situations react to the receipt of new information in a way that suggests they are following Bayes' rule, in other situations they appear to violate this rule; see, for instance, Robinson and Hastie [1985], Baratgin and Politzer [2007], Zhao et al. [2012], and Douven and Schupbach [2015a], [2015b]. Those violations could have been entirely unsystematic, of course, but Douven and Schupbach [2015a], [2015b] found that, in some contexts, they arise systematically because in accommodating new information people take into account explanatory considerations. This finding was not entirely surprising, given that explanatory considerations were already known to play various roles in cognition; see Chi et al. [1994], Sloman [1994], [1997], Lombrozo [2006], [2007], [2012], Legare, Wellman, and Gelman [2009], Legare

[2012], Williams and Lombrozo [2013], Legare and Lombrozo [2014], and Walker et al. [2014], among others. However, none of the previous studies had looked at the relationship between explanation and belief *change*.

That many formal epistemologists hesitate to consider such *explanatory* models of credence-redistribution is likely due to arguments asserting that *any* form of belief change at variance with Bayes' rule betokens irrationality. For anyone buying into such arguments, probabilistic versions of IBE may be of interest to psychology (where they may be studied alongside other types of irrational behavior), but they should hold no appeal to philosophers who focus on normative issues. Even with respect to the normative claim, however, the bad reputation of probabilistic versions of IBE may well be undeserved, as the motivating arguments themselves appear doubtful.

For instance, van Fraassen [1989] has made much of the fact that IBE in general will fail whenever the truth is not among the hypotheses under consideration. As explained in Schupbach [2014], however, that has no more significance than the fact that conjunction introduction will lead us to infer a false conclusion whenever one of the premises is false. Van Fraassen [1989] has also leveled Lewis' [1999] so-called dynamic Dutch book argument specifically against probabilistic versions of IBE. That argument purports to show that anyone whose belief changes are guided by such a version of IBE can be made to engage in series of bets the net payoff of which is necessarily negative. The argument is contentious, however (Douven [1999]), and even if it were sound, it is now generally recognized to concern practical rationality rather than epistemic rationality, which is the type of rationality at issue in a debate about we ought to change our graded beliefs. Finally, it has been argued that changing our graded beliefs by any other rule than Bayes' makes those graded beliefs less accurate than they would otherwise be (Rosenkrantz [1992]), but that turns out to be true only on one specific understanding of accuracy, where this understanding also appears to be of lesser epistemic importance (Douven [2013], [2016]).

References

Arló-Costa, H. [2006] "Review of *Tracking Truth: Knowledge, Evidence and Science*, by S. Roush (Oxford University Press)," *Notre Dame Philosophical Reviews* (available at `https://ndpr.nd.edu`).

Arló-Costa, H. and Parikh, R. [2006] "Tracking Truth: Knowledge and Conditionals in the Context of Branching Time," Unpublished.

Baratgin, J. and Politzer, G. [2007] "The Psychology of Dynamic Probability Judgment: Order Effect, Normative Theories, and Experimental Methodology," *Mind and Society* 6:53–66.

Bird, A. and Pettigrew, R. [2016] "Internalism, Externalism, and the KK Principle," Unpublished.

BonJour, L. [1985] *The Structure of Empirical Knowledge*, Cambridge, Mass.: Harvard University Press.

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

Boyd, R. [1984] "The Current Status of Scientific Realism," in J. Leplin (ed.) *Scientific Realism*, Berkeley CA: University of California Press, pp. 195–222.

Bradley, R. [2001] "Ramsey and the Measurement of Belief," in D. Corfield and J. Williamson (eds.) *Foundations of Bayesianism*, London: Kluwer Academic Publishers, pp. 263–290.

Briggs, R., Cariani, F., Easwaran, K., and Fitelson, B. [2014] "Individual Coherence and Group Coherence," in J. Lackey (ed.) *Essays in Collective Epistemology*, Oxford: Oxford University Press, pp. 215–239.

Brössel, P. and Eder, A.-M. A. [2014] "How to Resolve Doxastic Disagreement," *Synthese* 191:2359–2381.

Cevolani, G. [2014] "Truth Approximation, Belief Merging, and Peer Disagreement," *Synthese* 191:2383–2401.

Chandler, J. [2010] "The Lottery Paradox Generalized?" *British Journal for the Philosophy of Science* 61:667–679.

Chandler, J. [2013] "Acceptance, Aggregation and Scoring Rules," *Erkenntnis* 78:201–217.

Chi, M. T. H., de Leeuw, N., Chiu, M.-H., and LaVancher, C. [1994] "Eliciting Self-explanations Improves Understanding," *Cognitive Science* 18:439–477.

Christensen, D. [2004] *Putting Logic in its Place*, Oxford: Oxford University Press.

Dancy, J. [2003] "A Defense of Coherentism," in L. Pojman (ed.) *The Theory of Knowledge: Classical and Contemporary Readings*, Belmont, CA: Wadsworth, pp. 206–215.

Decock, L. B., Douven, I., and Sznajder, M. [2016] "A Geometric Principle of Indifference," *Journal of Applied Logic*, in press.

de Finetti, B. [1937/1964] "Foresight: Its Logical Laws, its Subjective Sources," in H. E. Kyburg and H. E. Smokler (eds.) *Studies in Subjective Probability*, New York: Wiley, pp. 55–118.

Dietrich, F. and List, C. [2010] "The Aggregation of Propositional Attitudes: Towards a General Theory," *Oxford Studies in Epistemology* 3:215–234.

Douven, I. [1999] "Inference to the Best Explanation Made Coherent," *Philosophy of Science* 66:S424–S435.

Douven, I. [2002] "A New Solution to the Paradoxes of Rational Acceptability," *British Journal for the Philosophy of Science* 53:391–410.

Douven, I. [2011] "Abduction," in E. N. Zalta (ed.) *Stanford Encyclopedia of Philosophy* (available at `http://plato.stanford.edu/archives/spr2011/entries/abduction/`).

Douven, I. [2013] "Inference to the Best Explanation, Dutch Books, and Inaccuracy Minimisation," *Philosophical Quarterly* 69:428–444.

Douven, I. [2016] "Inference to the Best Explanation: What Is It? And Why Should We Care?" in T. Poston and K. McCain (eds.) *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press, in press.

Douven, I. and Meijs, W. [2007a] “Measuring Coherence,” *Synthese* 156:405–425.

Douven, I. and Meijs, W. [2007b] “On the Alleged Impossibility of Coherence,” *Synthese* 157:347–360.

Douven, I. and Schupbach, J. N. [2015a] “The Role of Explanatory Considerations in Updating,” *Cognition* 142:299–311.

Douven, I. and Schupbach, J. N. [2015b] “Probabilistic Alternatives to Bayesianism: The Case of Explanationism,” *Frontiers in Psychology* 6:459, doi: 10.3389/fpsyg.2015.00459.

Douven, I. and Uffink, J. [2003] “The Preface Paradox Revisited,” *Erkenntnis* 59:389–420.

Douven, I. and Wenmackers, S. [2016] “Inference to the Best Explanation versus Bayes’ Rule in a Social Setting,” *British Journal for the Philosophy of Science*, in press.

Douven, I. and Williamson, T. [2006] “Generalizing the Lottery Paradox,” *British Journal for the Philosophy of Science* 57:755–779.

Earman, J. and Salmon, W. C. [1992] “The Confirmation of Scientific Hypotheses,” in M. H. Salmon, J. Earman, C. Glymour, and J. Lennox (eds.) *Introduction to the Philosophy of Science*, New York: Prentice-Hall, pp. 42–103.

Easwaran, K. [Forthcoming] “Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probabilities,” *Noûs*.

Elqayam, S. and Over, D. E. [2013] “New Paradigm Psychology of Reasoning,” *Thinking and Reasoning* 19:249–265.

Fitelson, B. [2003] “A Probabilistic Theory of Coherence,” *Analysis* 63:194-199.

Fitelson, B. and Jehle, D. [2009] “What is the ‘Equal Weight View’?” *Episteme* 6:280–293.

Foley, R. [1992] “The Epistemology of Belief and the Epistemology of Degrees of Belief,” *American Philosophical Quarterly* 29:111–124.

Fumerton, R. [1980] “Induction and Reasoning to the Best Explanation,” *Philosophy of Science* 47:589–600.

Fumerton, R. [1995] *Metaepistemology and Skepticism*, Boston: Rowman & Littlefield Publishers.

Glass, D. H. [2005] “Problems with Priors in Probabilistic Measures of Coherence,” *Erkenntnis* 63:375–85.

Goldman, A. I. [1986] *Epistemology and Cognition*, Cambridge MA: Harvard University Press.

Goldman, A. I. [1999] *Knowledge in a Social World*, Oxford: Oxford University Press.

Goldman, A. I. [2009] “Internalism, Externalism, and the Architecture of Justification,” *The Journal of Philosophy* 106(6):309–338.

Gwiazda, J. [2010] “Infinitism, Completability, and Computability: Reply to Peijnenburg,” *Mind* 119:1123–1124.

Haack, S. [1993] *Evidence and Inquiry*, Oxford: Blackwell.

Harman, G. [1965] “The Inference to the Best Explanation,” *Philosophical Review* 74:88–95.

Hartmann, S., Pigozzi, G., and Sprenger, J. [2010] "Reliable Methods of Judgment Aggregation," *Journal for Logic and Computation* 20:603–617.

Hawthorne, J., Landes, J., Wallmann, C., and Williamson, J. [2016] "The Principal Principle Implies the Principle of Indifference," *British Journal for the Philosophy of Science*, in press.

Hegselmann, R. and Krause, U. [2002] "Opinion Dynamics and Bounded Confidence: Models, Analysis, and Simulations," *Journal of Artificial Societies and Social Simulation* 5 (available at `http://jasss.soc.surrey.ac.uk/5/3/2.html`).

Hegselmann, R. and Krause, U. [2006] "Truth and Cognitive Division of Labor: First Steps towards a Computer Aided Social Epistemology," *Journal of Artificial Societies and Social Simulation* 9 (available at `http://jasss.soc.surrey.ac.uk/9/3/10.html`).

Hempel, C. G. [1962] "Deductive-Nomological vs. Statistical Explanation", in: H. Feigl and G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science* (Vol. 3), Minneapolis: University of Minnesota Press, pp. 98–169.

Henderson, L. [2014] "Bayesianism and Inference to the Best Explanation," *British Journal for the Philosophy of Science* 65(4):687–715.

Hintikka, J. [1962] *Knowledge and Belief*, Ithaca NY: Cornell University Press.

Howson, C. [2000] *Hume's Problem: Induction and the Justification of Belief*, Oxford: Clarendon Press.

Huemer, M. [2009] "Explanationist Aid for the Theory of Inductive Logic," *British Journal for the Philosophy of Science* 60:345–375.

Huemer, M. [2011] "Does Probability Theory Refute Coherentism?" *The Journal of Philosophy* 108(1):35–54.

Hume, D. [1748/1912] *An Enquiry Concerning Human Understanding*, Chicago: Open Court Publishing.

Jeffrey, R. [2004] *Subjective Probability: The Real Thing*, Cambridge: Cambridge University Press.

Joyce, J. M. [2005] "How Probabilities Reflect Evidence," *Philosophical Perspectives* 19:153–178.

Kahneman, D., Slovic, P., and Tversky, A. [1982] *Judgment under Uncertainty: Heuristics and Biases*, Cambridge: Cambridge University Press.

Kelly, K. T. [2014] "A Computational Learning Semantics for Inductive Empirical Knowledge," in A. Baltag and S. Smets (eds.) *Johan van Benthem on Logic and Information Dynamics*, Dordrecht: Springer, pp. 289–337.

Kelp, C. [2016] "Lotteries and Justification," *Synthese*, in press.

Keynes, J. M. [1921] *A Treatise of Probability*, London: Macmillan.

Klein, P. [1985] "The Virtues of Inconsistency," *Monist* 68:105–135.

Klein, P. [1998] "Foundationalism and the Infinite Regress of Reasons," *Philosophy and Phenomenological Research* 58(4):919–925.

Klein, P. and Warfield, T. A. [1994] "What Price Coherence?" *Analysis* 54(3):129–132.

Koehler, D. J. [1991] "Explanation, Imagination, and Confidence in Judgment," *Psychological Bulletin* 110:499–519.

Kornhauser, L. A. and Sager, L. G [1986] "Unpacking the Court," *Yale Law Journal* 96:82–117.

Kolmogorov, A. N. [1950] *Foundations of the Theory of Probability*, New York: Chelsea Publishing Company.

Kroedel, T. [2012] "The Lottery Paradox, Epistemic Justification and Permissibility," *Analysis* 72:57–60.

Kvart, I. [2006] "A Probabilistic Theory of Knowledge," *Philosophy and Phenomenological Research* 72(1):1–43.

Kyburg, H. [1961] *Probability and the Logic of Rational Belief*, Middletown CT: Wesleyan University Press.

Lam, B. [2011] "On the Rationality of Belief-Invariance in Light of Peer Disagreement," *Philosophical Review* 120:207–245.

Legare, C. H. [2012] "Exploring Explanation: Explaining Inconsistent Information Guides Hypothesis-testing Behavior in Young Children," *Child Development* 83:173–185.

Legare, C. H. and Lombrozo, T. [2014] "Selective Effects of Explanation on Learning During Early Childhood," *Journal of Experimental Child Psychology* 126:198–212.

Legare, C. H., Wellman, H. M., and Gelman, S. A. [2009] "Evidence for an Explanation Advantage in Naïve Biological Reasoning," *Cognitive Psychology* 58:177–194.

Leitgeb, H. [2014] "The Stability Theory of Belief," *Philosophical Review* 123:131–171.

Lewis, D. K. [1980] "A Subjectivist's Guide to Objective Chance," in R. C. Jeffrey (ed.) *Studies in Inductive Logic and Probability* (Vol. 2), Berkeley CA: University of California Press, pp. 83–132.

Lewis, D. K. [1999] *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press.

Levinstein, B. A. [2015] "With All Due Respect: The Macro-Epistemology of Disagreement," *Philosophers' Imprint* 15 (available at `http://hdl.handle.net/2027/spo.3521354.0015.013`).

Lipton, P. [2004] *Inference to the Best Explanation* (2nd ed.), London: Routledge.

Lin, H. and Kelly, K. T. [2012] A Geo-Logical Solution to the Lottery Paradox, with Applications to Conditional Logic," *Synthese* 186:531–575.

List, C. and Puppe, C. [2009] "Judgment Aggregation: A Survey," in C. List and C. Puppe (eds.) *Handbook of Rational and Social Choice*, Oxford: Oxford University Press.

Lombrozo, T. [2006] "The Structure and Function of Explanations," *Trends in Cognitive Sciences* 10:464–470.

Lombrozo, T. [2007] "Simplicity and Probability in Causal Explanation," *Cognitive Psychology* 55:232–257.

Lombrozo, T. [2012] "Explanation and Abductive Inference," in K. J. Holyoak and R. G. Morrison (eds.) *Oxford Handbook of Thinking and Reasoning*, Oxford: Oxford University Press, pp. 260–276.

Maher, P. [1993] *Betting on Theories*, Cambridge: Cambridge University Press.

Makinson, D. [1965] "The Paradox of the Preface," *Analysis* 25:205–207.

Mayo-Wilson, C. [2016] "(How) Is Scientific Knowledge Possible?" Unpublished.

McGrew, L. and McGrew, T. [2008] "Foundationalism, Probability, and Mutual Support," *Erkenntnis* 68(1):55–77.

McGrew, T. [1999] "How Foundationalists do Crossword Puzzles," *Philosophical Studies* 96:333–350.

McGrew, T. [2001] "Direct Inference and the Problem of Induction," *The Monist* 84(2):153–178.

McGrew, T. [2003] "Confirmation, Heuristics, and Explanatory Reasoning," *British Journal for the Philosophy of Science* 54:553-–567.

McGrew, T. and McGrew, L. [2000] "Foundationalism, Transitivity, and Confirmation," *Journal of Philosophical Research* 25:47–66.

Meacham, C. J. G. [2010] "Unravelling the Tangled Web: Continuity, Internalism, Non-Uniqueness and Self-Locating Beliefs," *Oxford Studies in Epistemology* 3:86–125.

Meyer, J.-J. and van der Hoek, W. [1995] *Epistemic Logic for AI and Computer Science*, Cambridge: Cambridge University Press.

Nelkin, D. [2000] "The Lottery Paradox, Knowledge, and Rationality," *Philosophical Review* 109:373–409.

Nozick, R. [1981] *Philosophical Explanations*, Cambridge, MA: Harvard University Press.

Oaksford, M. and Chater, N. [2007] *Bayesian Rationality*, Oxford: Oxford University Press.

Oaksford, M. and Chater, N. [2013] "Dynamic Inference and Everyday Conditional Reasoning in the New Paradigm," *Thinking & Reasoning* 19:346–379.

Okasha, S. [2013] "On a Flawed Argument Against the KK Principle," *Analysis* 73(1):80–86.

Olsson, E. J. [2002] "What is the Problem of Coherence and Truth?," *Journal of Philosophy* 94:246–272.

Olsson, E. J. [2008] "Knowledge, Truth, and Bullshit: Reflections on Frankfurt," *Midwest Studies in Philosophy* 32:94–110.

Olsson, E. J. [2011] "A Simulation Approach to Veritistic Social Epistemology," *Episteme* 8:127–143.

Olsson, E. J. and Vallinder, A. [2013] "Norms of Assertion and Communication in Social Networks," *Synthese* 190:2557–2571.

Over, D. E. [2009] "New Paradigm Psychology of Reasoning," *Thinking and Reasoning* 15:431–438.

Peijnenburg, J. [2007] "Infinitism Regained," *Mind* 116:597–602.

Peijnenburg, J. [2010] "Ineffectual Foundations: Reply to Gwiazda," *Mind* 119:1125–1133.

Peijnenburg, J. and Atkinson, D. [2008] "Probabilistic Justification and the Regress Problem," *Studia Logica* 89:333–341.

Pettigrew, R. [2015] "What Chance–Credence Norms Should Not Be," *Noûs* 49:177–196.

Pollock, J. [1990] *Nomic Probability and the Foundations of Induction*, Oxford: Oxford University Press.

Priest, G. [1998] "What Is so Bad about Contradictions?" *Journal of Philosophy* 95:410–426.

Ramsey, F. P. [1926/1990] "Truth and Probability," in D. H. Mellor (ed.) *Philosophical Papers*, Cambridge: Cambridge University Press, pp. 52–94.

Robinson, L. B. and Hastie, R. [1985] "Revision of Beliefs when a Hypothesis Is Eliminated from Consideration," *Journal of Experimental Psychology: Human Perception and Performance* 11(4):443–456.

Rosenkrantz, R. D. [1992] "The Justification of Induction," *Philosophy of Science* 59:527–539.

Roush, S. [2005] *Tracking Truth: Knowledge, Evidence, and Science*, Oxford: Oxford University Press.

Salmon, W. C. [2001] "Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation," in G. Hon and S. S. Rakover (eds.) *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer, pp. 61–91.

Schupbach, J. N. [2008] "On the Alleged Impossibility of Bayesian Coherentism," *Philosophical Studies* 141(3):323–331.

Schupbach, J. N. [2011a] "New Hope for Shogenji's Coherence Measure," *British Journal for the Philosophy of Science* 62(1):125–142.

Schupbach, J. N. [2011b] *Studies in the Logic of Explanatory Power*, PhD thesis, University of Pittsburgh.

Schupbach, J. N. [2014] "Is the Bad Lot Objection Just Misguided?" *Erkenntnis* 79:55–64.

Schupbach, J. N. [2016] "Inference to the Best Explanation, Cleaned Up and Made Respectable," in T. Poston and K. McCain (eds.) *Best Explanations: New Essays on Inference to the Best Explanation*, Oxford: Oxford University Press, in press.

Shogenji, T. [1999] "Is Coherence Truth-Conducive?" *Analysis* 59(4):338–345.

Shogenji, T. [2003] "A Condition for Transitivity in Probabilistic Support," *The British Journal for the Philosophy of Science* 54(4):613–616.

Siebel, M. and Schippers, M. [2015] "Inconsistency as a Touchstone for Coherence Measures," *Theoria. An International Journal for Theory, History and Foundations of Science* 30(1):11–41.

Sloman, S. A. [1994] "When Explanations Compete: The Role of Explanatory Coherence on Judgments of Likelihood," *Cognition* 52:1–21.

Sloman, S. A. [1997] "Explanatory Coherence and the Induction of Properties," *Thinking and Reasoning* 3:81–110.

Smith, M. [2010] "A Generalised Lottery Paradox for Infinite Probability Spaces," *British Journal for the Philosophy of Science* 61:821–831.

Sprenger, J., Martini, C., and Hartmann, S. [2009] "Consensual Decision-Making Among Epistemic Peers," *Episteme* 6:110–129.

Staley, K. and Cobb, A. [2010] "Internalist and Externalist Aspects of Justification in Scientific Inquiry," *Synthese* 182(3):475–492.

Tentori, K., Crupi, V., Bonini, N. and Osherson, D. [2007] "Comparison of Confirmation Measures," *Cognition* 103:107–119.

Thagard, P. [1989] "Explanatory Coherence," *Behavioral and Brain Sciences* 12:435–502.

Thagard, P. [2000] *Coherence in Thought and Action*, Cambridge, Mass: MIT Press.

Vallinder, A. and Olsson, E. J. [2014] "Trust and the Value of Overconfidence: A Bayesian Perspective on Social Network Communication," *Synthese* 191:1991–2007.

van Fraassen, B. C. [1989] *Laws and Symmetry*, Oxford: Clarendon Press.

Vogel, J. [1990] "Cartesian Skepticism and Inference to the Best Explanation," *Journal of Philosophy* 87:658–666.

Walker, C. M., Lombrozo, T., Legare, C., and Gopnik, A. [2014] "Explaining Prompts Children to Privilege Inductively Rich Properties," *Cognition* 133:343–357.

Wenmackers, S. [2013] "Ultralarge Lotteries: Analyzing the Lottery Paradox Using Non-Standard Analysis," *Journal of Applied Logic* 11:452–467.

Weisberg, J. [2009] "Locating IBE in the Bayesian Framework," *Synthese* 167:125–143.

Williams, J. J. and Lombrozo, T. [2013] "Explanation and Prior Knowledge Interact to Guide Learning," *Cognitive Psychology* 66:55–84.

Williamson, T. [2000] *Knowledge and its Limits*, Oxford: Oxford University Press.

Zalabardo, J. L. [2012] *Skepticism and Reliable Belief*, Oxford: Oxford University Press.

Zhao, J., Crupi, V., Tentori, K., Fitelson, B., and Osherson, D. [2012] "Updating: Learning versus supposing," *Cognition* 124:373–378.