



Research

Cite this article: Pothos EM, Lewandowsky S, Basieva I, Barque-Duran A, Tapper K, Khrennikov A. 2021 Information overload for (bounded) rational agents. *Proc. R. Soc. B* **288**: 20202957.

<https://doi.org/10.1098/rspb.2020.2957>

Received: 27 November 2020

Accepted: 11 January 2021

Subject Category:

Behaviour

Subject Areas:

behaviour

Keywords:

Bayesian inference, disagreement, entrenchment, rationality, decision-making

Author for correspondence:

Emmanuel M. Pothos

e-mail: emmanuel.pothos.1@city.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5287575>.

Information overload for (bounded) rational agents

Emmanuel M. Pothos¹, Stephan Lewandowsky², Irina Basieva¹, Albert Barque-Duran³, Katy Tapper¹ and Andrei Khrennikov⁴

¹Department of Psychology, City, University of London, London EC1V 0HB, UK

²School of Psychological Science, University of Bristol, Bristol BS8 1TU, UK

³Department of Computer Science, Universitat de Lleida, Carrer de Jaume II, 67, 25001 Lleida, Spain

⁴International Center for Mathematical Modeling in Physics and Cognitive Science, Linnaeus University, Universitetplatsen 1, 351 95 Växjö, Sweden

EMP, 0000-0003-1919-387X; AK, 0000-0002-9857-0938

Bayesian inference offers an optimal means of processing environmental information and so an advantage in natural selection. We consider the apparent, recent trend in increasing dysfunctional disagreement in, for example, political debate. This is puzzling because Bayesian inference benefits from powerful convergence theorems, precluding dysfunctional disagreement. Information overload is a plausible factor limiting the applicability of full Bayesian inference, but what is the link with dysfunctional disagreement? Individuals striving to be Bayesian-rational, but challenged by information overload, might simplify by using Bayesian networks or the separation of questions into knowledge partitions, the latter formalized with quantum probability theory. We demonstrate the massive simplification afforded by either approach, but also show how they contribute to dysfunctional disagreement.

1. Background

Truthiness is tearing apart our country ... It used to be, everyone was entitled to their own opinion, but not their own facts. But that's not the case anymore

—Stephen Colbert, January 2006

Living organisms depend on the optimal processing of environmental information, for example, regarding foraging, mate selection or the assessment of predation risks. Environmental information is typically uncertain, and so has to be processed probabilistically. The established standard for probabilistic inference is Bayesian probability theory [1] (we will refer to it as just Bayesian theory or occasionally full Bayesian theory, for emphasis). Bayesian theory provides a set of mutually coherent principles for probabilistic reasoning on uncertain premises. Bayesian theory benefits from powerful normative arguments, such as the Dutch book theorem, which shows that Bayesian probabilities will never lead to inconsistencies, such as certain loss in a combination of gambles [1]. Accordingly, Bayesian reasoning is often characterized as rational. There is an immense body of work successfully validating Bayesian models of human cognition [2–4]; these models are not universally successful, but they are successful enough to allow confidence that humans can be sometimes rational in the Bayesian sense.

Moreover, for non-human animals, it has been argued that Bayesian inference confers a natural selection advantage [5,6] and there have been simulations of how natural selection enables the computation of Bayesian priors across generations [7] or other aspects of Bayesian behaviour [8] (the first step in probabilistic inference is the determination of priors, that is, the assumptions regarding the probabilities of relevant events prior to any new information). Evidence for animal behaviour consistent with Bayesian inference has been observed in, for example, foraging [9] or mating [10] (overview in [11]). The requirement of optimality in animal behaviour is often grounded in Bayesian

terms, even acknowledging that Bayesian consistency may be focused on particular environments or circumstances [8,12].

However, for both humans and non-human animals, there have been inconsistencies between Bayesian principles and behaviour. For humans, some evocative examples have been produced by the influential work of Tversky & Kahneman. For example, Tversky & Kahneman [13] described a hypothetical person, Linda, as outgoing, concerned with equality, and intellectually restless. Naive participants considered it more likely that Linda is a bank teller and a feminist, than just a bank teller. Such conjunction fallacies challenge Bayesian intuition at a fundamental level; it is like judging that it is more likely to rain and snow in December, than just snow. Interestingly, analogous fallacies appear in animal behaviour too. For example, rhesus macaques can show ambiguity aversion [14] and pigeons sometimes show the less is more effect, whereby a desirable food plus a less desirable food is perceived less appealing than the desirable food alone [15].

As Valone [11, p. 257] noted, 'Greater attention needs to be devoted to understanding when and when not to expect Bayesian updating and to determine the limits of Bayesian updating in animals'. The exact point applies to human behaviour too. Here, we pursue a novel perspective to the emergence of non-Bayesian behaviour in humans, motivated by the apparent increase in dysfunctional disagreement in, e.g. modern political debate. We call dysfunctional disagreement when it appears impossible for two parties to converge, regardless of iterations and evidence. Our analysis is not restricted to political debate, but it is easier to develop the argument this way.

The evidence for increasing dysfunctional disagreement and deterioration in the quality of political debate is strong. For example, consider the emergence of 'truthiness', as in Colbert's quote above (based on his satirical show), which can be defined as 'truth that comes from the gut, not books' [16]; consider also the increasing dissemination of 'fake news' [17] and their ability to set the political agenda [18]; and the intense polarization surrounding recent political events (e.g. the Brexit referendum vote in the UK). Kahan [19, p. 1] offers an evocative quote: 'Never have human societies known so much about mitigating the dangers they face but agreed so little about what they collectively know'.

It is tempting to consider these points unsurprising, because there is a staggering range of factors contributing to disagreement, particularly when people rely on false information [20]. Disagreement may arise due to emotional influences. Emotion can overwhelm objective information [21] or bias the activated information [22]. Some theorists suggest that all reasoning is motivated [23], so that discourse is guided just by insistence on a particular position. Differences in values can result in persistent disagreement [24]. For example, conflicts between a refutation message for a prior position and valued self-conceptions may lead people to become more entrenched [25]. There are several related biases. For example, the disconfirmation bias is scepticism for premises incongruent with one's beliefs [26]. The 'mybias' is collecting information and assessing evidence in a way biased in favour of a person's beliefs [27]. Mybias is especially problematic in information-rich societies, since plurality and freedom of expression mean that one can find supporting opinions for any position. For example, Del Vicario *et al.* [28] argued that information related to distinct narratives generates homogeneous, polarized communities

on Facebook. Such echo chambers could embody contradictory perspectives between them [29] and lead to distorted pictures regarding consensus.

We focus on individuals striving to be (i) as Bayesian as possible, (ii) up to date with the relevant information and (iii) willing to put aside their egos in the interest of resolving disagreements constructively. We call such individuals *well-meaning*, and also suggest that they can set aside unmovable personal values (i.e. we need not worry about disagreement from values [24]). Such well-meaning individuals should be able to avoid most of the 'standard' sources of disagreement. For example, in dual decision routes, analytic versus intuitive components [30] correspond to thoughtful versus spontaneous cognition. Bayesian inference might be predominantly localized in the analytic route; but the relative balance between different routes is partly under conscious control, depending on effort, time, etc. Or Bayesian inference might be reflected in the intuitive route, with non-Bayesian behaviour arising from limitations from working memory or language when accessing the basis of intuitive judgements [31]. But it should be possible to reduce such limitations, with effort. Also, decision biases might be avoidable with the adoption of behavioural rules [32]; it is known that emotions can be monitored and their impact on behaviour limited [33], etc.

Here is the paradox: more people are educated than ever before in history, there is more insight regarding decision biases, we have better understanding of the importance of the common good, and access to information has never been easier. All these factors should increase our capacity for Bayesian cognition. At the very least, we can assume that the proportion of well-meaning individuals in society has not changed, maybe even increased (would we not like to consider ourselves as well-meaning?). So, why does it appear that increasingly there is dysfunctional disagreement surrounding many current debates?

We suggest that even for well-meaning individuals, information overload challenges our capacity for Bayesian thought, in a way that leads to dysfunctional disagreement. It is easiest to make our case in relation to political debate, but the ideas are general. First, we ask whether there is increasing information overload in political debate. The case is straightforward. One cause of information overload is the multiplicity of media and ways to disseminate information in modern society. Practically every second, the Internet, television, mobile phones, etc. pump out massive amounts of news, comments on the news and comments on the comments. Another cause is that in a technologically advanced society, some debates are complex, for example, because they relate to technological innovations that cannot be easily comprehended in lay terms. Access to information has never been easier and we enjoy unprecedented benefits from technological advancement, yet these factors contribute to massive information overload.

Second, we consider whether information overload might contribute to dysfunctional disagreement. There are indications that this is the case [34]. Allenby & Sarewitz [35] suggest that the technological complexity of modern society is such that informed decisions are beyond the scope of comprehension for the majority of us. John [36] suggests that scientists best serve society by *relaxing* the maxims of transparency and openness—not because openness and transparency are undesirable, but because too much information may

damage public trust in science, because the public's folk philosophy of science is at odds with the actual workings of science. There is clearly a pessimistic view concerning whether people can deal with the information complexity in modern political debates [37,38].

We develop a precise link between information overload and non-Bayesian inference and consider the implications for dysfunctional disagreement, even for well-meaning individuals. It is interesting that animal behaviour researchers have also considered whether information overload (environmental complexity) might challenge Bayesian processes [39].

2. Outline of methods

We consider two well-meaning individuals, Alice and Bob, debating a question and examine their capacity for avoiding dysfunctional disagreement, under conditions of information overload. Convergence means agreement on at least the probabilities for question outcomes, noting that in complex debates, it is rarely the case there are uncontested observations, even for good faith actors. We quantify information overload in terms of the number of ancillary questions that inform our decision on a key question. For example, suppose Alice is interested in the Brexit question. She could inform her eventual decision on Brexit by considering questions such as 'Will Brexit be good for the economy?', 'Will Brexit be good for employment rights?', etc., noting that each of these questions could be further broken down. There is information overload when the number of these ancillary questions increases beyond a 'practical' point.

Can well-meaning individuals agree to disagree? Bounded rationality is the form of rationality which emerges when the resources of the reasoning agent are insufficient for full rationality. So, what are forms of bounded rationality under conditions of information overload and the implications for dysfunctional disagreement?

3. Disagreement and Bayesian rationality

Consider well-meaning Alice and Bob debating a complex political question and assume they share their questions and outcomes. They then use their respective information to define a probability distribution and update their beliefs as rational Bayesian agents. Is it possible for Alice and Bob to dysfunctionally disagree? Suppose Alice and Bob have different information regarding a Brexit question, but share priors and have common knowledge of each other's posteriors (posteriors are the updated probabilities, once some new information has been received). Then Aumann's [40] theorem guarantees that Alice and Bob's posteriors will be the same, that is, two rational agents will eventually converge. Moreover, this convergence can be achieved with a reasonable amount of effort [41]. The requirement of common priors may appear stringent; however, it can be replaced by milder ones [42]. Even without common priors, Bayesian Alice and Bob willing to share information must eventually converge. The Bernstein–von Mises theorem guarantees that Bayesian updating will converge posteriors (as long as there is no 'zero priors' trap [43]). Finally, some of these results depend on honest exchange of information. For well-meaning Alice and Bob, this should be straightforward, assuming they can agree on acceptable error bounds.

Overall, well-meaning Bayesian Alice and Bob committed to full Bayesian inference cannot agree to disagree [41,42].

How practical is it for Alice and Bob to be fully Bayesian under conditions of information overload? The essential idea is this (see also electronic supplementary material, S1). Consider a finite set Ω of all possible elementary events (the most specific events which can occur) and all possible subsets, including the null set \emptyset and Ω itself. This set theoretic representation of events is appropriate if each event is either true or not true.¹ We can perform logical operations on these subsets, union, intersection and complementation, which correspond to the familiar operations of conjunction, disjunction and negation. The requirement that each of these operations produces a subset of Ω enables an algebra over the space of subsets, which is a Boolean algebra (because the operations obey commutativity, associativity and distributivity). We can then define a probability measure over these subsets, which is a map from the space of subsets to the real number interval $[0, 1]$, with normalization 1 for Ω .

Consider Alice confronted with questions $A, B, C, D \dots$, each of which can have possible outcomes $A_1 \dots A_m, B_1 \dots B_m$, etc. Each block of question outcomes generates its own Boolean algebra, $\beta(A), \beta(B), \dots$. Before Alice can engage with probabilistic reasoning for a question, she first needs to construct these individual Boolean algebras, which involves a process of specifying conjunctions, disjunctions and negations of outcomes. But, for a Bayesian Alice confronted with questions, A, B, \dots, F , it is insufficient to have $\beta(A), \beta(B) \dots \beta(F)$. For a consistent *joint probability distribution* across any combination of question outcomes, she also needs to construct a bigger Boolean algebra $\beta(A, B, \dots, F)$, which integrates the algebras for the individual questions in a consistent way. This larger algebra requires knowledge of conjunctions and disjunctions for all the individual question outcomes A_i, \dots, F_j , belonging to the different algebras $\beta(A), \beta(B), \dots \beta(F)$.

The problem of intractability of full Bayesian representations is well known, cf. the idea of magic sets in Artificial Intelligence [44]. We illustrate it in the case of debating, for example, Brexit and ancillary questions, such as whether Brexit might be good for the economy, labour laws, etc. If we had nine binary ancillary questions, then the elementary events would be enumerated as

1. Brexit_{yes}, X1_{yes} ... X9_{yes}
2. Brexit_{yes}, X1_{yes} ... X9_{no}
- ...
1024. Brexit_{no}, X1_{no} ... X9_{no}

Given these $2^{10} = 1024$ elementary events, we can evaluate any more elaborate question, for example, a conjunction involving some question outcomes versus others, such as $\text{Prob}(X1_{\text{yes}} \& X2_{\text{yes}} \text{ or } X3_{\text{yes}} \& X5_{\text{no}})$. But, the immense expressive power of Bayesian theory comes with the price of requiring knowledge of the joint probability distribution—here, the probabilities of all 1024 elementary events. The more questions we have, the more complex the joint probability distribution and so any probabilistic inference. As the number of questions n and outcomes per question k increase, the number of terms in the joint probability distribution increase as k^n .

To quantify complexity, we adopt an information-theoretic coding scheme and compute *information costs* [45,46]

(electronic supplementary material, S1). The coding cost of D numbers can be specified by dividing the relevant number range into D bins and assigning each number to one bin, which requires $\log_2 D$ bits for each number for a total of $D \log_2 D$ bits. This is intuitive because if the D numbers were uniformly distributed, we would have enough bins to just make them discriminable (if $D = 100$, these statements are equivalent to representing the numbers with two decimal places; electronic supplementary material, S2). Therefore, the information cost for representing probabilistic information for n questions with k outcomes each is $(k^n - 1) \log_2(k^n - 1)$ bits, approximated as $k^n \log_2 k^n$.

Information overload clearly undermines full Bayesian inference. Consider a person living in an isolated community a hundred years ago. He would be confronted with a fairly limited range of questions, each of which would be affected by relatively few events. So, it would be undemanding to create a Boolean algebra of all questions, including conjunctions, disjunctions, etc. Today, especially in political debate, we are confronted with questions of immense complexity. Consider Alice faced with the Brexit dilemma. There are hundreds of questions relevant to resolving the dilemma, across several categories, for example, relating to finance, immigration, security, and so on. Alice does not have the time or resources (mental or otherwise) to create a full Boolean algebra for all questions and their outcomes.

When confronted with a complex probability distribution, a powerful approach is sampling algorithms, such as Markov chain Monte Carlo (MCMC) methods [3,47,48]. An MCMC method will approximate Bayesian computations, by employing samples from the probability distribution, instead of the full distribution. Such samples are often selected to favour more probable parts of the distribution and depending on the similarity of the parts already selected. However, in the present case, sampling approximations will not help: when faced with problems of increasing complexity, sampling from the full distribution will delay, but not avoid, the exponential explosion of probability terms.

4. Bayesian networks

The first approach we consider for mitigating the problems of complex distributions is Bayesian networks (e.g. [49]). Suppose we recognize that in many cases, questions will be independent of each other, so that e.g. $\text{Prob}(A|B) = \text{Prob}(A)$ or conditionally independent so that, for example, $\text{Prob}(A \& B|X) = \text{Prob}(A|X) \cdot \text{Prob}(B|X)$. Clearly, such an approach has simplifying potential, since a complex conditional probability $\text{Prob}(A|X_1, X_2, X_3, X_4 \dots)$ might be easily computable as, for example, $\text{Prob}(A|X_1)$. The way to formalize assumptions about conditional independence is Bayesian networks. Bayesian networks represent (acyclic) probabilistic relations between a set of variables, such that each variable is a node and causal relations are represented as directed edges. The simplifying potential of Bayesian networks rests with their Markov property: without causal dependencies, there are no conditional dependencies. So, simplification depends on the causal structure. Note, there is extensive evidence for the psychological plausibility of Bayesian networks [50,51], even if it is unclear whether they suffice for a cognitive theory of causality [52]. Presently,

we are only concerned with the way the local Markov property can simplify probabilistic information.

If Alice and Bob are overwhelmed by the complexity of their representations, they could use Bayesian networks as a simplifying tactic. But it is unlikely they will develop similar causal structures for their representations, as these would depend on their experience, education, background, etc. Bayesian networks Alice and Bob with different causal structures mean that the powerful classical convergence theorems (Aumann's theorem; the Bernstein–von Mises theorem) no longer hold. Alice and Bob could now find themselves in a state of dysfunctional disagreement, even though they are fully rational, given their representations (which correspond to different assumptions regarding causal structure). Alice and Bob could seek convergence by communicating their causal structure, but such knowledge is often hard to articulate. Note, there have been attempts to explain dysfunctional disagreement with Bayesian networks with hidden nodes corresponding to, for example, attitudes which prevent convergence [53,54]. The present point is related, but instead concerns the inevitable incidental differences in causal structures.

To estimate the complexity of probabilistic inference with Bayesian networks, consider classical Alice contemplating six binary questions related to the Brexit question. Without the Markov property, the probability distribution for a particular combination of question outcomes would look like $\text{Prob}(X1_{\text{yes}}, X2_{\text{yes}}, X3_{\text{yes}}, Y1_{\text{yes}}, Y2_{\text{yes}}, Y3_{\text{yes}}, \text{Brexit}_{\text{yes}}) = \text{Prob}(X1_{\text{yes}}|X2_{\text{yes}}, X3_{\text{yes}}, Y1_{\text{yes}}, Y2_{\text{yes}}, Y3_{\text{yes}}, \text{Brexit}_{\text{yes}}) \times \text{Prob}(X2_{\text{yes}}|X3_{\text{yes}}, Y1_{\text{yes}}, Y2_{\text{yes}}, Y3_{\text{yes}}, \text{Brexit}_{\text{yes}}) \dots \text{Prob}(\text{Brexit}_{\text{yes}})$. The Markov property allows us to assume certain questions to be independent. For example, regarding $\text{Prob}(A|X, Y)$, we may be able to write $\text{Prob}(A|X, Y) = \text{Prob}(A|X)$. Suppose that Alice employing a Bayesian network assumes partial conditional independence, so that conditionalizations depend on m variables. Then, we would write, if $m = 2$, $\text{Prob}(X1_{\text{yes}}, X2_{\text{yes}}, X3_{\text{yes}}, Y1_{\text{yes}}, Y2_{\text{yes}}, Y3_{\text{yes}}, \text{Brexit}_{\text{yes}}) = \text{Prob}(X1_{\text{yes}}|A_{\text{yes}}, B_{\text{yes}}) \cdot \text{Prob}(X2_{\text{yes}}|C_{\text{yes}}, D_{\text{yes}}) \dots$, where A, B are two questions on which $X1$ depends etc. As long as $m \ll n$, each term requires k^m probabilities (ignoring '-1'), for a total of approximately $n \cdot k^m$ probabilities [55]. The associated coding complexity for the joint probability distribution given a particular Bayesian network is $n \cdot k^m \log_2(n \cdot k^m)$ bits. We also need the information cost of specifying a Bayesian network, and can show that overall the information cost for probabilistic information encoded using a Bayesian network is $(n \cdot k^m) \log_2(n \cdot k^m) + n \left[\log_2 \binom{n-1}{m} + \log_2 n \right]$ (electronic supplementary material, S2).

(a) Quantum probability theory: disagreement

We call quantum theory the probability rules from quantum mechanics, without the physics. Behaviours that appear classically erroneous can sometimes have simple explanations in quantum theory, which motivates the psychological plausibility of such models [56–58].

Informally, quantum theory is just like Bayesian theory for subsets of questions (compatible sets, see below), but across these subsets, apparent classical errors can arise. These incompatible sets are like knowledge partitions, segments of knowledge such that within each segment, but not across segments, reasoning is rational. Knowledge partitions

can emerge as a simplifying strategy in complex problems [59,60]. For example, when learning an association between two variables based on a complex function, a natural approach is to learn the association in smaller ranges, but in a way that the corresponding parts are not integrated with each other. Well-meaning Alice dealing with Brexit might try to be rational for specific subsets of questions, but without trying to integrate the Boolean algebra for one theme with another. For example, if Alice works in the financial sector, she may be able to create a full Boolean structure regarding the financial implications from Brexit and so be rational for such questions. At the same time, Alice is so busy with the construction of this finance Boolean algebra, that she does not have time to do the same for other Brexit questions, e.g. relating to security. Arguably, this is what we are seeing in modern society: individuals highly knowledgeable and rational in specific areas but who, when asked to consider questions across other areas, may be challenged and even produce inconsistent beliefs.

In quantum theory, instead of a set Ω of elementary events, we have a Hilbert space H , such that each vector in H corresponds to an elementary event (a Hilbert space is essentially a complex vector space with a scalar product). Question outcomes correspond to subspaces in H ; each subspace is associated with a projector P (which ‘lays’ down a vector onto a subspace); in psychological theory, the mental state is represented by a normalized vector in H ; probabilities are computed by projecting the state vector onto subspaces and squaring the length of the projections. Different partitions in H are defined by sets of basis vectors. For example, in a standard coordinate space, we might have three basis vectors along the x, y, z directions. Basis sets are not unique. If we apply the same rotation to each of our current vectors x, y, z , we will end up with a new set of basis vectors x', y', z' . Two sets of basis vectors can be related to each other using a generalized kind of rotation.

Projectors can be compatible, in which case we have a Boolean algebra exactly as in the classical case, or incompatible, when the Boolean algebra structure breaks down. That is, considering sets $A, B, C \dots$ of projectors, such that within each set projectors are compatible, but across incompatible sets, one *cannot* combine Boolean algebras $\beta(A), \beta(B) \dots$ into one large Boolean algebra. Each event in this larger structure is no longer either true or not true (before measurement) and distributivity is no longer obeyed. Instead, we have a partial Boolean algebra, which is a collection of Boolean algebras pasted together, so that where any two Boolean algebras overlap, their operations agree. Conjunctions and disjunctions preserve their Boolean features only within the same Boolean algebra. Conjunctions of incompatible questions have a sequential form and $\text{Prob}(P_A \wedge \text{then } P_B) \neq \text{Prob}(P_B \wedge \text{then } P_A)$. Also, a definite answer for a question can create uncertainty for other incompatible ones.

Quantum theory can simplify probabilistic inference with incompatibility, which allows Alice to squeeze information about, say, 100 questions (which, even if binary, will require a classical space of 2^{100} dimensions) into a space of, say, 10 dimensions. If quantum Alice organizes her large set of Brexit questions into incompatible themes, each theme corresponds to a basis set in the same small dimensionality space and the representation of new themes need only involve a change of basis, instead of enlargement of the original space. However, incompatibility contributes to dysfunctional disagreement.

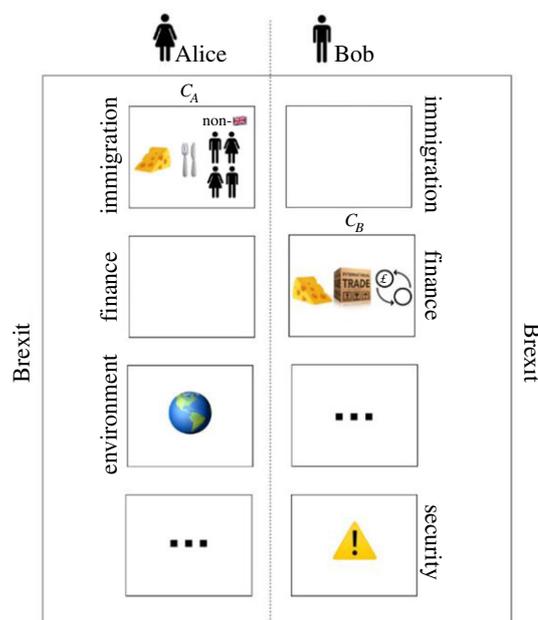


Figure 1. Alice and Bob are interested in whether Brexit may increase the price of imported cheese, C . Alice considers C with questions related to immigration, while Bob considers C with finance questions. As a result, Alice and Bob develop meanings for the C question which are different, even though they think they are considering the same question. (Online version in colour.)

One implication of incompatibility is that quantum Alice is more likely to display (classical) fallacies, which may undermine her arguments. Incompatibility has been linked with conjunction and disjunction fallacies [61], question order effects [62], violations of normative constraints in causal reasoning [51] and disjunction effects [63]. Moreover, incompatibility leads to contextuality in meaning. If quantum Alice and Bob have different partial Boolean algebras, they may think they are talking about the same question, have the same data and fail to agree, because they are talking about different questions (figure 1). Such ideas resemble proposals in social psychology about how earlier questions can activate thoughts or perspectives for later ones [64]. Contextuality arises in quantum theory because the meaning of question A is determined by considering the set of questions compatible with A (and some of these questions might be incompatible with each other) and because the meaning of question A may be affected by considering prior questions incompatible with A .

Contextuality contributes to dysfunctional disagreement. First, quantum Alice and Bob are no longer aided by Aumann’s theorem [65]. Common knowledge in the quantum case is not equivalent to common knowledge in the classical case, because the former lacks conjunctions. Additionally, questions incompatible with common knowledge will produce interference terms so that Alice and Bob will not update probabilities consistently with each other. Second, collective decision-making typically benefits from communal knowledge effects, such as the community of knowledge effect, wisdom-of-the-crowds and Condorcet’s jury theorem. Such effects are not specific to Bayesian inference, but they are consistent with it. However, all three are undermined by contextuality. Regarding community of knowledge, Sloman & Fernbach [66] argued that in a complex world, we increasingly benefit from each other’s expertise and sometimes, as a result, overestimate our own knowledge (a knowledge

illusion). The wisdom-of-the-crowds effect is the proposal that an averaged judgement across observers can be more accurate than most individual judgements, assuming primarily independence of observations and that individual estimates are normally distributed around the correct outcome [67]. Finally, the Condorcet jury theorem shows that a majority decision (e.g. in a jury) is increasingly likely to be correct, as we add voters whose (individual) probability that they are correct is just over 0.5. Regarding community of knowledge and wisdom of the crowds, if Alice and Bob are debating contextual question A , then Alice may be thinking of A_X and Bob of A_Y , where X, Y indicate differing meanings. This casts doubt on the rationality of putting Alice's and Bob's intuitions together. Such problems are likely to be accentuated, because employing a partial Boolean algebra may lead to overconfidence (electronic supplementary material, S3).

(b) Quantum theory: coding costs

Within a single partition, we have a classical probability distribution for the corresponding questions, encoded in the mental state vector. We need to specify the mental state for one partition and the way partitions relate to each other; the latter is encoded in transformation operators called unitary. So, the information cost for probabilistic inference for quantum Alice depends on three elements, the mental state vector for one partition, unitary operators and the cost of allocating questions to partitions. The mental state vector and unitary operators are specified in terms of parameters which are real numbers. Regarding information costs, we follow from the above approach to assume that F real parameters (assumed in a certain range) can be approximately specified using $F \log_2 F$ bits.

Label the dimensionality of each partition as N . The mental state vector in N dimensions has $N-1$ real parameters corresponding to amplitudes and $N-1$ real parameters for the phases. This is because the N amplitudes are constrained by the normalization condition and, regarding the N phases, the quantum state is the same up to an overall phase factor. The corresponding information cost is $2 \cdot (N-1) \log_2(N-1)$, which can be approximated as $2 \cdot N \log_2 N$. What is N ? Suppose c partitions are employed and that all partitions have the same number of questions. Then, in each partition, we have n/c questions, k outcomes each, so that $N = k^{n/c}$. The overall information cost involves additional terms, for how information in one partition relates to information in other partitions. This cost is $2 \cdot k^{n/c} \log_2 k^{n/c} + (c-1) \cdot \frac{4n}{c} \log_2 \frac{4n}{c} + \log_2 \frac{n!}{\left[\left(\frac{n}{c}\right)!\right]^{c-1}} + \log_2 \frac{(c-1)! n^c}{c^{c-1}}$ (electronic supplementary material, S2). Note, the dimensionality of quantum Alice's probability space turns out to be only $N = k^{n/c}$, which seems like a huge saving compared to Bayesian Alice for whom $N = k^n$; but this simplification is partly offset by the complexity of specifying partition relations.

5. Comparisons

A well-meaning Alice overwhelmed by the complexity of her joint probability distribution might seek to simplify the representations either by employing Bayesian networks or dividing her questions into (incompatible) partitions. For the latter two schemes, the critical parameters are,

respectively, m (the average number of questions each one question depends on) and c (the number of partitions). Both parameters concern the extent of dependence of questions among themselves and, specifically, the length of conditional probabilities (electronic supplementary material, S2). Regarding m , this interpretation follows directly from the definition of a Bayesian network, while in the quantum case, classical conditionalization occurs only within knowledge partitions. Therefore, it is natural to set $n/c = m$ or $c = n/m$.

We provide indicative estimates regarding the simplification from Bayesian networks and quantum theory relative to Bayesian theory, varying question numbers from 5 to 15 and question outcomes from 2 to 4, figure 2. The vertical axis shows the information cost for scheme A (e.g. Bayesian theory) minus B (e.g. Bayesian networks). Recall, lower information costs are more advantageous, so that when $A - B \gg 0$, then B is superior to A. In all cases, probabilistic reasoning with either Bayesian networks or quantum theory affords overwhelming simplification relative to Bayesian theory. This is a demonstration of the essential point that information overload will drive even well-meaning Alice to make representational approximations, putatively employing Bayesian networks or knowledge partitions.

We also observe a marginal advantage of quantum theory over Bayesian networks, though this conclusion is sensitive to the complexity of the relation between partitions. Overall, the quantum approach to simplification seems advantageous, thus providing a strong expectation of dysfunctional disagreement due to incompatibility and partitions.

6. Concluding comments

We considered how dysfunctional disagreement can arise for well-meaning individuals, because of information overload. The notion of being well-meaning is primarily underwritten by an assumption of rational cognition, in the Bayesian sense. There is a strong consensus that Bayesian rationality is achievable to some extent [1–4]. Our aim has been to understand how information overload can challenge full Bayesian rationality, how Bayesian networks and quantum theory offer flavours of limited or local Bayesian rationality, and the implications for dysfunctional disagreement.

Regarding dysfunctional disagreement, a full Bayesian would quickly find it impossible to build the required Boolean algebra, for complex problems. Alice can simplify with Bayesian networks, truncating her probability distributions with assumptions about the causal structure between her questions. Alice and Bob may find themselves failing to converge if their Bayesian networks are different; Aumann's [40] and the Bernstein–von Mises theorems no longer hold. Alternatively, Alice can simplify using knowledge partitions [59] dividing her questions into sets, such that within each knowledge partition, she is fully Bayesian, but across partitions apparent errors arise. With knowledge partitions, Aumann's and the Bernstein–von Mises theorems also no longer hold and, in addition, the resulting contextuality challenges the community of knowledge effect [66], wisdom-of-the-crowds effect [67] and Condorcet's jury theorem.

Is it possible for Bayesian networks or quantum Alice and Bob to converge? In the former case, they need to share their causal structure. However, it seems unlikely this would occur, because we are often unaware of the causal dependencies impacting on inference. In the latter case, Alice and Bob need to share their partitions (and information on how partitions

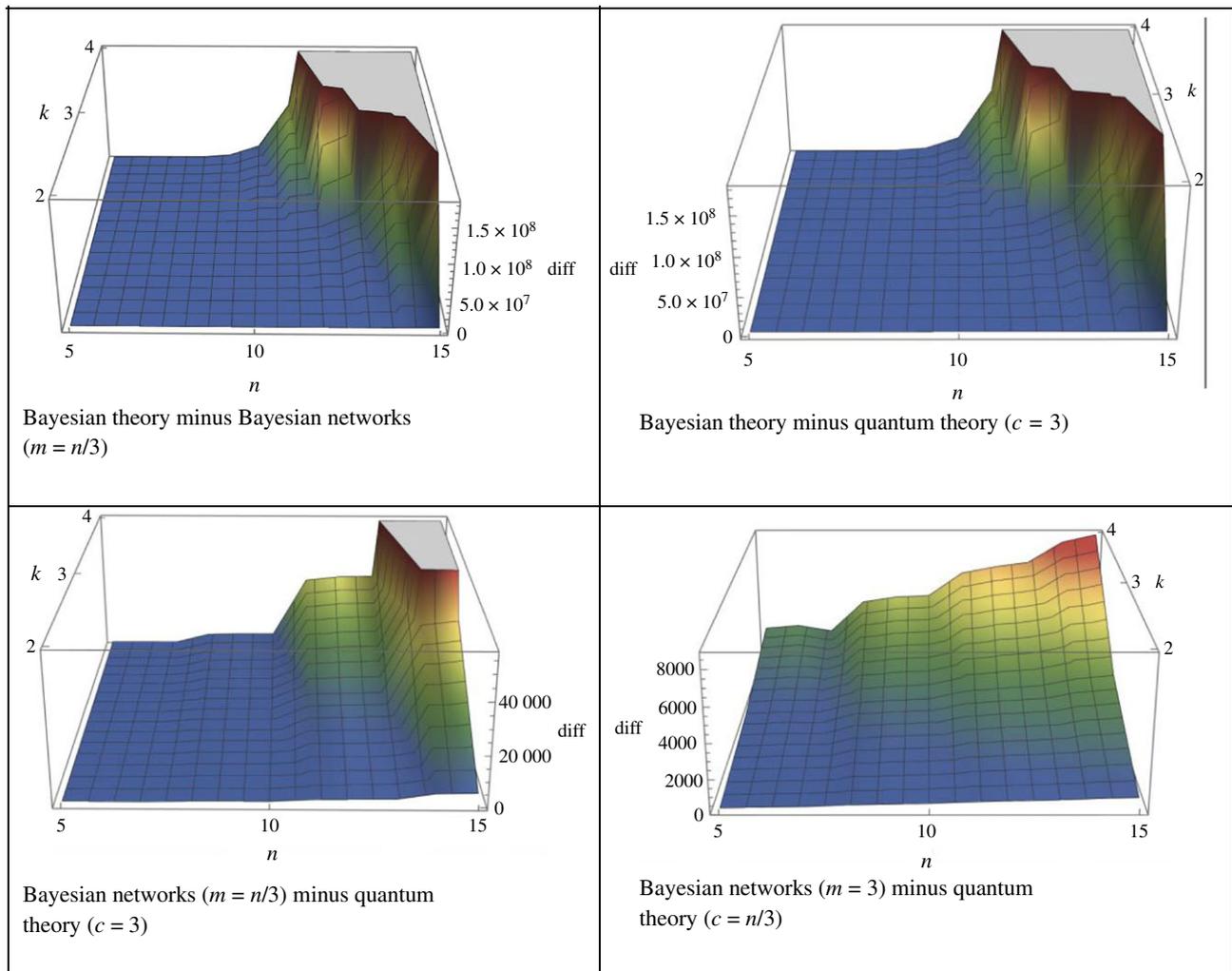


Figure 2. We plot information cost given one scheme minus information cost given another scheme, labelled Diff (in bits). The superior scheme has lower information cost. Horizontal axes represent the number of questions (n) and outcomes per question (k); complexity increases with both n and k . Note, $m = 3$ for Bayesian networks translates to three questions per knowledge partition in QPT. (Online version in colour.)

relate to each other), and in addition be careful to respond to a question in the same context (figure 1). We agree with Lissack [68], who argued that truthiness can be reduced if Alice and Bob 'Try to see things from my viewpoint'. However, we think quantum Alice and Bob will not engage with such a process, because contextuality is not recognized in probabilistic inference.

Our focus has been dysfunctional disagreement, because this is an under-researched topic and because the link with information overload is intuitive. More generally, there have been long research traditions concerning the way complexity undermines Bayesian rationality. The present framework can shed light into other instances of behaviour apparently problematic from a full Bayesian perspective, because of complexity, bearing in mind that there will be behaviours outside any probabilistic framework. For example, the emergence of some conjunction fallacies, as in the Linda example [13], could be traced to lack of familiarity with partition combinations. It is possible that we have a local partition for professions and one for personal characteristics, like feminism, without making the effort to combine them together. Conversely, the less-is-more effect in animal behaviour [15] seems harder to understand as complexity-driven bounded rationality.

In closing, to the long list of factors contributing to dysfunctional disagreement, we add differences in causal structure and contextuality, from information overload. A surprising implication is that more information or nuanced

perspectives may exacerbate disagreement by further encouraging truncated probability distributions or incompatible representations as simplifying tactics. For some important modern debates, such as Brexit, it may seem that we have forgotten how to evaluate arguments using easily verifiable facts, but increasing information may not help or may indeed be harmful [36–38]. A precise understanding of the impact of information overload, as we have offered, will hopefully contribute to mitigating interventions.

Data accessibility. Computer code available at: <https://osf.io/h9tjk>.

Authors' contributions. A.K. and I.B. initially conceived of the idea. E.M.P. provided the initial versions of the models and analyses. E.M.P. and A.K. provided the initial drafts of this work. All authors worked to develop the ideas and the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the Office of Naval Research Global (grant number N62909-19-1-2000).

Acknowledgements. Thanks to Adam Sanborn for MCMC advice and three anonymous reviewers.

Endnote

¹On any Boolean algebra, it is possible to define a truth function taking values 'true' or 'not true'. On a partial Boolean algebra, such a truth function cannot be introduced.

1. Oaksford M, Chater N. 2009 Précis of Bayesian rationality: the probabilistic approach to human reasoning. *Behav. Brain Sci.* **32**, 69–120. (doi:10.1017/S0140525X09000284)
2. Griffiths TL, Chater N, Kemp C, Perfors A, Tenenbaum JB. 2010 Probabilistic models of cognition: exploring representations and inductive biases. *Trends Cogn. Sci.* **14**, 357–364. (doi:10.1016/j.tics.2010.05.004)
3. Sanborn AN, Griffiths TL, Navarro DJ. 2010 Rational approximations to rational models: alternative algorithms for category learning. *Psychol. Rev.* **117**, 1144–1167. (doi:10.1037/a0020511)
4. Tenenbaum JB, Kemp C, Griffiths TL, Goodman N. 2011 How to grow a mind: statistics, structure, and abstraction. *Science* **331**, 1279–1285. (doi:10.1126/science.1192788)
5. Oaten A. 1977 Optimal foraging in patches: a case for stochasticity. *Theor. Popul. Biol.* **12**, 263–285. (doi:10.1016/0040-5809(77)90046-6)
6. Iwasa Y, Higashi M, Yamamura N. 1981 Prey distribution as a factor determining the choice of optimal foraging strategy. *Am. Nat.* **117**, 710–723. (doi:10.1086/283754)
7. Ramirez JC, Marshall JAR. 2017 Can natural selection encode Bayesian priors? *J. Theor. Biol.* **426**, 57–66. (doi:10.1016/j.jtbi.2017.05.017)
8. McNamara JM, Green RF, Olsson O. 2006 Bayes' theorem and its application in animal behaviour. *Oikos* **112**, 243–251. (doi:10.1111/j.0030-1299.2006.14228.x)
9. Valone TJ, Giraldeau LA. 1993 Patch estimation by group foragers: what information is used? *Anim. Behav.* **45**, 721–728. (doi:10.1006/anbe.1993.1086)
10. Luttbeg B, Warner RR. 1999 Reproductive decision-making by female peacock wrasses: flexible versus fixed behavioral rules in a variable environment. *Behav. Ecol.* **10**, 666–674. (doi:10.1093/beheco/10.6.666)
11. Valone TJ. 2006 Are animals capable of Bayesian updating? An empirical review. *Oikos* **112**, 252–259. (doi:10.1111/j.0030-1299.2006.13465.x)
12. Trimmer PC, Houston AI, Marshall JAR, Mendl MT, Paul ES, McNamara JM. 2011 Decision-making under uncertainty: biases and Bayesians. *Anim. Cogn.* **14**, 465–476. (doi:10.1007/s10071-011-0387-4)
13. Tversky A, Kahneman D. 1983 Extensional versus intuitive reasoning: the conjunctive fallacy in probability judgment. *Psychol. Rev.* **90**, 293–315. (doi:10.1037/0033-295X.90.4.293)
14. Hayden BY, Heilbronner SR, Platt ML. 2010 Ambiguity aversion in rhesus macaques. *Front. Neurosci.* **4**, 1–7. (doi:10.3389/fnins.2010.00166)
15. Zentall TR, Laude JR, Case JP, Daniels CW. 2014 Less means more for pigeons but not always. *Psychon. Bull. Rev.* **21**, 1623–1628. (doi:10.3758/s13423-014-0626-1)
16. Meyer D. 2006 The truth of truthiness. *CBS News*.
17. Allcott H, Gentzkow M. 2017 Social media and fake news in the 2016 election. *J. Econ. Perspect.* **31**, 211–236. (doi:10.1257/jep.31.2.211)
18. Vargo CJ, Guo L, Amazeen MA. 2018 The agenda-setting power of fake news: a big data analysis of the online media landscape from 2014 to 2016. *New Media Soc.* **20**, 2028–2049. (doi:10.1177/1461444817712086)
19. Kahan DM. 2015 What is the 'science of science communication'? *J. Sci. Commun.* **14**, 1–2. (doi:10.22323/2.14030404)
20. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. 2012 Misinformation and its correction: continued influence and successful debiasing. *Psychol. Sci. Public Interest* **13**, 106–131. (doi:10.1177/1529100612451018)
21. Wilson RS, Arvai JL. 2006 When less is more: how affect influences preferences when comparing low and high risk options. *J. Risk Res.* **9**, 165–178. (doi:10.1080/13669870500419503)
22. Bower GH. 1981 Mood and memory. *Am. Psychol.* **36**, 129–148. (doi:10.1037/0003-066X.36.2.129)
23. Kunda Z. 1990 The case for motivated reasoning. *Psychol. Bull.* **108**, 480–498. (doi:10.1037/0033-2909.108.3.480)
24. Schwartz SH. 1992 Universals in the content and structure of values: theory and empirical tests in 20 countries. In *Advances in experimental social psychology*, vol. 25 (ed. M Zanna), pp. 1–65. New York, NY: Academic Press.
25. Trevors GJ, Muis KR, Pekrun R, Sinatra GM, Winne PH. 2016 Identity and epistemic emotions during knowledge revision: a potential account for the backfire effect. *Discourse Process.* **53**, 339–370. (doi:10.1080/0163853X.2015.1136507)
26. Taber CS, Lodge M. 2006 Motivated skepticism in the evaluation of political beliefs. *Am. J. Polit. Sci.* **50**, 755–769. (doi:10.1111/j.1540-5907.2006.00214.x)
27. Stanovich KE, West RF, Toplak ME. 2013 Myside bias, rational thinking, and intelligence. *Curr. Dir. Psychol. Sci.* **22**, 259–264. (doi:10.1177/0963721413480174)
28. Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, Stanley HE, Quattrociocchi W. 2016 The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559. (doi:10.1073/pnas.1517441113)
29. Pariser E. 2011 *The filter bubble: what the internet is hiding from you*. London, UK: Penguin.
30. Kahneman D. 2001 *Thinking fast and slow*. London, UK: Penguin.
31. Oaksford M, Hall S. 2016 On the source of human irrationality. *Trends Cogn. Sci.* **20**, 336–344. (doi:10.1016/j.tics.2016.03.002)
32. Gollwitzer PM, Sheeran P. 2006 Implementation intentions and goal achievement: a meta-analysis of effects and processes. *Adv. Exp. Soc. Psychol.* **38**, 69–119.
33. Chambers R, Gullone E, Allen NB. 2009 Mindful emotion regulation: an integrative review. *Clin. Psychol. Rev.* **29**, 560–572. (doi:10.1016/j.cpr.2009.06.005)
34. Lissack M. 2016 Don't be addicted: the oft-overlooked dangers of simplification. *J. Des. Econ. Innov.* **2**, 29–45. (doi:10.1016/j.sheji.2016.05.001)
35. Allenby BR, Sarewitz D. 2011 We've made a world we cannot control. *New Scientist*, 14 May, 28–29.
36. John S. 2018 Epistemic trust and the ethics of science communication: against transparency, openness, sincerity and honesty. *Soc. Epistemol.* **32**, 75–87. (doi:10.1080/02691728.2017.1410864)
37. Hills TT. 2019 The dark side of information proliferation. *Perspect. Psychol. Sci.* **14**, 323–330. (doi:10.1177/1745691618803647)
38. Sunstein CR. 2020 *Too much information: understanding what you don't want to know*. Cambridge, MA: MIT Press.
39. Valone TJ. 1992 Information for patch assessment: a field investigation with black-chinned hummingbirds. *Behav. Ecol.* **3**, 211–222. (doi:10.1093/beheco/3.3.211)
40. Aumann RJ. 1976 Agreeing to disagree. *Annu. Stat.* **4**, 1236–1239. (doi:10.1214/aos/1176343654)
41. Aaronson S. 2005 The complexity of agreement. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pp. 634–643. Baltimore, MD: ACM.
42. Hanson R. 2006 Uncommon priors require origin disputes. *Theory Decis.* **61**, 319–328. (doi:10.1007/s11238-006-9004-4)
43. Basieva I, Pothos EM, Trueblood J, Khrennikov A, Busemeyer JR. 2017 Quantum probability updating from zero priors (by-passing Cromwell's rule). *J. Math. Psychol.* **77**, 58–69. (doi:10.1016/j.jmp.2016.08.005)
44. Bancilhon F, Maier D, Sagiv Y, Ullman JD. 1986 Magic sets and other strange ways to implement logic programs. In *PODS '86: Proceedings of the fifth ACM SIGACT-SIGMOD symposium on Principles of database systems*, pp. 1–16. Baltimore, MD: ACM.
45. Oaksford M, Chater N. 1994 A rational analysis of the selection task as optimal data selection. *Psychol. Rev.* **101**, 608–631. (doi:10.1037/0033-295X.101.4.608)
46. Pothos EM, Chater N. 2002 A simplicity principle in unsupervised human categorization. *Cogn. Sci.* **26**, 303–343. (doi:10.1207/s15516709cog2603_6)
47. Lloyd K, Sanborn A, Leslie D, Lewandowsky S. 2019 Why higher working memory capacity may help you learn: sampling, search, and degrees of approximation. *Cogn. Sci.* **43**, e12805. (doi:10.1111/cogs.12805)
48. Vul E, Goodman N, Griffiths TL, Tenenbaum JB. 2014 One and done? Optimal decisions from very few samples. *Cogn. Sci.* **38**, 599–637. (doi:10.1111/cogs.12101)
49. Gershman SJ, Vul E, Tenenbaum JB. 2012 Multistability and perceptual inference. *Neural Comput.* **24**, 1–24. (doi:10.1162/NECO_a_00226)
50. Rehder B. 2014 Independence and dependence in human causal reasoning. *Cognit. Psychol.* **72**, 54–107. (doi:10.1016/j.cogpsych.2014.02.002)

51. Trueblood JS, Yearsley JM, Pothos EM. 2017 A quantum probability framework for human probabilistic inference. *J. Exp. Psychol. Gen.* **146**, 1307–1341. (doi:10.1037/xge0000326)
52. Goldvarg E, Johnson-Laird PM. 2001 Naïve causality: a mental model theory of causal meaning and reasoning. *Cogn. Sci.* **25**, 565–610. (doi:10.1207/s15516709cog2504_3)
53. Jern A, Chang KMK, Kemp C. 2014 Belief polarization is not always irrational. *Psychol. Rev.* **121**, 206–224. (doi:10.1037/a0035941)
54. Cook J, Lewandowsky S. 2016 Rational irrationality: modeling climate change belief polarization using Bayesian networks. *Topics Cogn. Sci.* **8**, 160–179. (doi:10.1111/tops.12186)
55. Murphy K. 1998 A brief introduction to graphical models and Bayesian networks. See <https://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html>.
56. Busemeyer JR, Bruza P. 2011 *Quantum models of cognition and decision making*. Cambridge, UK: Cambridge University Press.
57. Haven E, Khrennikov A. 2013 *Quantum social science*. Cambridge, UK: Cambridge University Press.
58. Pothos EM, Busemeyer JR. 2013 Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* **36**, 255–327. (doi:10.1017/S0140525X12001525)
59. Lewandowsky S, Kalish M, Ngang SK. 2002 Simplified learning in complex situations: knowledge partitioning in function learning. *J. Exp. Psychol. Gen.* **131**, 163–193. (doi:10.1037/0096-3445.131.2.163)
60. Kalish ML, Lewandowsky S, Kruschke JK. 2004 Population of linear experts: knowledge partitioning and function learning. *Psychol. Rev.* **111**, 1072–1099. (doi:10.1037/0033-295X.111.4.1072)
61. Busemeyer JR, Pothos EM, Franco R, Trueblood JS. 2011 A quantum theoretical explanation for probability judgment ‘errors’. *Psychol. Rev.* **118**, 193–218. (doi:10.1037/a0022542)
62. Wang Z, Solloway T, Shiffrin RM, Busemeyer JR. 2014 Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl Acad. Sci. USA* **111**, 9431–9436. (doi:10.1073/pnas.1407756111)
63. Pothos EM, Busemeyer JR. 2009 A quantum probability explanation for violations of ‘rational’ decision theory. *Proc. R. Soc. B* **276**, 2171–2178. (doi:10.1098/rspb.2009.0121)
64. Schwarz N. 2007 Attitude construction: evaluation in context. *Soc. Cogn.* **25**, 638–656. (doi:10.1521/soco.2007.25.5.638)
65. Khrennikov A, Basieva I. 2014 Possibility to agree on disagree from quantum information and decision making. *J. Math. Psychol.* **62**, 1–15. (doi:10.1016/j.jmp.2014.09.003)
66. Sloman S, Fernbach P. 2017 *The knowledge illusion: the myth of individual thought and the power of collective wisdom*. London, UK: Macmillan.
67. Surowiecki J. 2005 *The wisdom of crowds*. New York, NY: Anchor Books.
68. Lissack M. 2017 Explain the mess we are in: how did we get Trump, Brexit, dissonance, uncertainty and worse? See https://www.youtube.com/watch?v=Wqb2WYy_Rw.