Article
What Is Rational and Irrational in Human Decision Making

Emmanuel M. Pothos *, Oliver J. Waddup, Prince Kouassi and James M. Yearsley

Department of Psychology, University of London, London, EC1V 0HB, UK; Oliver.Waddup@city.ac.uk (O.J.W.); Prince.Kouassi@city.ac.uk (P.K.); James.Yearsley@city.ac.uk (J.M.Y.)
* Correspondence: emmanuel.pothos.1@city.ac.uk; Tel.: +44-(0)20-7040-0267

Abstract: There has been a growing trend to develop cognitive models based on the mathematics of quantum theory. A common theme in the motivation of such models has been findings which apparently challenge the applicability of classical formalisms, specifically ones based on classical probability theory. Classical probability theory has had a singularly important place in cognitive theory, because of its (in general) descriptive success but, more importantly, because in decision situations with low, equivalent stakes it offers a multiply justified normative standard. Quantum cognitive models have had a degree of descriptive success and proponents of such models have argued that they reveal new intuitions or insights regarding decisions in uncertain situations. However, can quantum cognitive models further benefit from normative justifications analogous to those for classical probability models? If the answer is yes, how can we determine the rational status of a decision, which may be consistent with quantum theory, but inconsistent with classical probability theory? In this paper, we review the proposal from Pothos, Busemeyer, Shiffrin, and Yearsley (2017), that quantum decision models benefit from normative justification based on the Dutch Book Theorem, in exactly the same way as models based on classical probability theory.

Keywords: classical probability theory; quantum probability theory; Dutch Book Theorem; question order effects; conjunction fallacy

1. Introduction: Decision Fallacies

Research in fallacies in decision making has produced some of the most evocative empirical results in behavioural sciences and attracted huge publicity and recognition (e.g., Nobel prizes for Kahneman and Thaler, both in economics). Part of the reason why this research has been so influential is because the corresponding results appear so surprising, given the established standards for rationality. Regarding the latter, the predominant framework for rational decision making in behavioural sciences has been classical probability theory. The justification for the rational status of classical probability theory can be established through multiple routes [1–3]. A widely considered one is the Dutch Book Theorem [4]. The Dutch Book Theorem has been appealing to behavioural scientists because it offers a perspective on rationality which is purely operational: rational assignment of probabilities to events needs to follow a simple coherence requirement, specifically, that probabilistic assignment to gambles has to be in such a way that the person is protected from certain loss regardless of how events turn out to be. The Dutch Book Theorem shows that if assignment of probabilities to events follows the basic axioms of classical probability theory, then there is no combination of corresponding gambles that will result in a sure loss. The Dutch Book Theorem offers a rational justification for decision models based on classical probability theory, as long as the stakes are small and equivalent. These ideas can break down if loss or risk aversion impact on decision making [5,6]. In addition to this strong rational foundation, classical probability theory is very intuitive. In a well-known quote from Laplace, we read " . . . [classical probability theory] is nothing but common sense reduced to calculation" [7] (p.313). This tenet of intuitiveness is of key relevance for behavioural sciences since, after all, it is human intuition (in probabilistic inference) that we are aiming to model.

In behavioural sciences, formal justification (the Dutch Book Theorem), descriptive success, and basic intuition have all contributed to make classical probability theory the dominant framework for decision theory. Whenever an empirical result at odds with classical prescription is identified, the most immediate approach is to explore ways to reconcile it with classical probability theory, then seek explanations based on alternative frameworks. The term 'decision fallacies' exactly refers to findings which are considered hard to reconcile with classical probability theory. A decision fallacy is not just about identifying a result which is inconsistent with classical probability theory. Rather, to characterise a decision as a fallacy it has to be the case that human observers experience persistence in the non-classical intuition, so that even when the classically correct result is explained, they cannot reject the non-classical intuition [8]. Therefore, what are some examples of results so classically problematic that they have challenged explanations based on classical probability theory?

Question order effects are particularly compelling, not least because in some cases they have been demonstrated in contexts with huge applied significance. Consider two sequentially presented binary questions, A and B, so that Prob(A) is the probability to respond to the A question with a yes and analogously for Prob(B). Therefore, if we encounter question A and then question B, the probability of responding with a yes to both questions should be given by Prob(A)Prob(B | A), noting that since the B question follows A, our answer to this question is informed by our prior answer to A. However, Prob(A)Prob(B | A) = Prob(A&B) = Prob(B&A) = Prob(B)Prob(A | B). Such an expression is trivially simple and shows that with two sequentially presented questions there should be no order effects. Responses should be identical whether the questions are presented one way or another. In some cases, human observers do indeed respect the classical constraint and respond consistently to question pairs regardless of orders. However, this is not always the case.

In a famous study, Moore et al. [9] examined participants' responses to a question about Clinton's honesty followed by a question about Gore's honesty versus the same two questions in the reverse order. When the Clinton question followed the Gore one, the averaged response rate for honesty was as high as 57%. When the Clinton question was first, the rate was only 50%. What is most shocking about this result is that it was observed in the context of a Gallup poll, large sample survey. Gallup poll outcomes often have a significant role in decision making, e.g., in relation to voting preferences for ambivalent individuals. Therefore, to see a difference of almost 20% as a result of a simple question order change provides a bleak perspective for the reliability of poll results and their capacity to provide measures of public attitudes which are unbiased or resistant to manipulation. One response to such a result is that poll responding might be reflexive or non-analytic [10,11] and so in such cases apparent classical fallacies simply reflect decision makers' lack of engagement with the task. However, analogous order effects are observed in cases of medical diagnosis with participants as medical professionals [12] and in mock jury decision making tasks [13–15]. For such cases, it would be hard to argue that classical decision errors arise from lack of sufficient consideration or attention of the corresponding questions.

Another fallacy challenging classical intuition concerns the law of total probability, the seemingly obvious constraint that, for example, for binary questions we require that Prob(A) = Prob(A & B) + Prob(A & ~B). Consider Shafir and Tversky's [16]'s variant of a prisoner's dilemma task. Participants are asked to make a one shot (no follow up decisions) decision of whether to defect or cooperate, based on a payoff matrix which depends both on the participant's decision and the decision of a hypothetical participant. The payoff matrix is set up so that the mutual payoff across both participants (i.e., both the real and the hypothetical one) is greatest when they both cooperate. However, the personal payoff for the real participant is greatest when he/she defects and the hypothetical one cooperates and vice versa. That is, prisoner's dilemma games test the tension between selfish interest and altruistic sentiments [17]. Shafir and Tversky's [16]'s design involved standard trials, the response of the hypothetical participant would not be known, as well as trials that

were labelled as bonus ones, for which the real participant was given the response of the hypothetical one prior to his/her own response. Perhaps unsurprisingly given the one-shot nature of the decisions, when participants were told the hypothetical one would cooperate, they decided to defect (since this is the highest payoff option); equally, when they were told that the hypothetical one would defect, they decided to defect too (since why would they cooperate given defection of the other party). What was surprising was that in the unknown case many participants switched their decision and decided to cooperate. That is, Shafir and Tversky [16] observed a situation where Prob(cooperate, unknown) > Prob(cooperate & opponent defect) + Prob(cooperate & opponent cooperate). With a such a paradigm, a classically oriented commentator might argue that the intricacies of altruism might ultimately confound consistency with probabilistic principles—that is, that there are additional decision principles at work over and above probabilistic inference. Ref. [16] reported a similar result for a gambling paradigm (see also [18]). In such cases, violations of the law of total probability in a gambling paradigm could be related to loss aversion. However, consider also the paradigm of Townsend et al. [19]. Participants always had to decide whether to attack or withdraw, in response to being presented with a face– faces would look more or less threatening and ostensibly this was the intended basis for participants' decisions. In some trials, the decision to attack or withdraw would be offered right after the presentation of the face. In other trials, participants would first have to categorize the face as good or bad. When participants categorized the face as bad, predictably there was a relatively (compared to the other cases) high probability to attack, Prob(Bad & Attack) = 0.52. Additionally, fairly predictably, when the face was categorized as good, the probability to attack was low, Prob(Good & Attack) = 0.07. What is astonishing is that when this categorization step was omitted, participants reversed their behaviour and were much more likely to attack, Prob(Attack, without categorization) = 0.69.

Perhaps the most evocative decision fallacy, and the last example in this brief overview, is the conjunction fallacy [20]. Consider a hypothetical person, Linda (arguably the most famous hypothetical person in decision sciences), who is described very much as a feminist and not at all as a bank teller. Participants receive the information about Linda and are then asked to rank order statements about her, from more to less likely ones. The key statements are whether Linda is a bank teller and a feminist. Surprisingly, participants' responses indicate that Prob(F&BT) > Prob(BT), violating the constraint from classical probability theory that a conjunction can never be more probable than a marginal. It could be argued that naïve observers can be excused from making such an error, because it is less natural to deal with probabilities expressed as subjective degrees of belief–arguably, humans have evolved to make probabilistic judgments based on relative frequencies. Ref. [21] offered a variant of the Linda problem that makes the set-theoretic structure of probabilities crystal clear. Participants are asked to evaluate the relative probabilities that a Scandinavian person has blond hair vs. blond hair and blue eyes. Results still revealed that Prob(blue eyes & blond hair) > Prob(blond hair). As a first impression, it is hard to accept this finding as anything other than a complete paradox. One can imagine lining up a group of Scandinavian individuals. How is it possible to have more individuals in the more restrictive category than the less restrictive one? This finding is the focus of our consideration of the rational status of a corresponding quantum approach. However, before this, we consider whether there are ways to salvage classical explanations of results such as the ones in this section.

2. Classical Accounts for Decision Fallacies?

There are ways to reconcile even the most apparently problematic decision fallacy with classical probability theory, not least because the key empirical results in decision sciences are often based on simpler paradigms with few data degrees of freedom. Take, for example, question order effects. Clearly, classically we cannot just write Prob(A)Prob(B|A) = ~Prob(B)Prob(A|B) (where the symbol = ~indicates lack of equality). However, we can write Prob(A)Prob(B|A & context1) = ~Prob(B)Prob(A|B & context2). The crucial

step here is conditionalizing on the context1 and context2 variables, which are taken to represent the varying frames of mind which are generated by considering prior questions. Specifically, in considering first question A, we can suppose that a particular frame of mind or context is generated, which impacts on how we approach question B–and vice versa. These ideas closely resonate with theory in social psychology, according to which prior questions do alter our mental state in a way that affects subsequent ones [22]. The problem with this account is that currently there are no principles in classical probability cognitive models which could allow us to set the context1 and context2 variables a priori. Therefore, just the presence of these variables allows us to cover any empirical result equally well, whether it is Prob(A)Prob(B | A & context1) > Prob(B)Prob(A | B & context2), Prob(A)Prob(B | A & context1) < Prob(B)Prob(A | B & context2), or Prob(A)Prob(B | A & context1) = Prob(B)Prob(A | B & context2). Such a model is devoid of explanatory value and, indeed, is not advocated in the literature.

A similar classical argument can be offered for violations of the law of total probability. Conditionalizations can be introduced which allow disparity between the marginal and the conjunctions. However, as above, lacking an a priori justification for how this conditionalizing variables can be set, it is difficult to consider the corresponding approach as a viable explanation.

The conjunction fallacy is a finding for which there has been particular effort to salvage a classical account. When non-specialists are confronted with the conjunction fallacy, the most immediate response is that participants do not understand the statements exactly as given and, instead, 'fill in' additional information. The idea that linguistic statements are augmented with additional information is well established [23]. Specific to the conjunction fallacy, it is possible that participants understand the marginal BT as BT & ~F, so that the corresponding comparison is not between Prob(BT) vs. Prob (BT & F) but rather between Prob(BT & ~F) vs. Prob(BT & F) [24]. However, Tentori et al. [21] (see also [25]), together with the A&B and A statements, also included A & ~B statements, so as to help participants discriminate between conjunctions and marginals–there is still a residual conjunction fallacy. In fact, across a large number of manipulations and changes, including using probabilities based on frequencies instead of subjective degrees of belief, there is a very persistent residual conjunction fallacy [26].

There have also been proposals of alternative frameworks for accommodating the conjunction fallacy, which are still based on classical probability theory. Authors in [27] employed a confirmation account to explain the conjunction fallacy. The idea here is not to evaluate Prob(BT & F | Linda) vs. Prob(BT | Linda), but rather Prob(F | BT & Linda) > Prob(F | BT); the latter clearly seems plausible. The objective of the model was to dissociate the emergence of a conjunction fallacy with one possibility being particularly likely and they presented some results supporting this novel empirical prediction. However, there is a price to pay in that conjunction is assessed with a function which is no longer a simple conjunction. As a result, it is difficult to extend this approach to disjunction errors, order effects, and other related fallacies [28]. There has also been some effort to understand conjunction fallacies as based on classical probabilities, but which are noisy [29,30]. These ideas assume that probabilities are based on a mental sampling process. As samples are generated, e.g., from memory or mental simulation, errors arise, which can then lead to a conjunction fallacy. There are two problems with such accounts. First, for a conjunction fallacy to be observed, one requires that the sampling error for conjunctions is higher than for marginals and that the sample process for marginals is independent from the one for conjunctions. Especially the latter is an unrealistic assumption, since why would we sample for, e.g., A and then separately for A & B? Second, such accounts either offer post hoc accounts of order effects [29] or no account at all [30], offering at best a limited account of fallacies in decision making.

It is such difficulties which have contributed to a desire to explore alternative formal frameworks for probabilistic inference. Of course, classical probability theory is hardly unique. In physics, it has been well-known that there is in fact an infinite hierarchy of

probabilistic systems, varying in the complexity of interference terms that are involved [31]. Therefore, why restrict cognitive modelling to just classical probability theory? Even if classical probability theory were to prove the most suitable framework for cognitive modelling (and the fallacies cast some doubt on this possibility), it would be amiss to not consider any alternatives. As a final note, in this brief introduction we focussed on classical probability theory models and their variants. There have been many other attempts to accommodate decision fallacies, notably based on heuristics and biases [32]. Arguments have been made to justify heuristics in decision making as rational, on the basis of optimality considerations [33,34]. These are very interesting approaches which have enhanced our understanding of decision making. Regarding rationality specifically, the case is harder to make, since considerations of optimality make it harder to separate between behaviours which can be considered rational vs. not.

3. Quantum Decision Models

Quantum probability theory, as employed by behavioural scientists, simply concerns the basic probability rules of quantum theory without any of the physics. Therefore, for example, Planck's constant is meaningless in behavioural sciences, as well as any of the commonly employed quantities in physics, such as position, momentum, energy, etc. It is important to note that a classical brain is assumed [35]: quantum principles, if at all relevant, are considered as potential descriptors of some aspects of behaviour; that is, whatever, the exact neurophysiological processes in the brain, we can examine whether there is structure in behaviour which can be characterised by quantum principles. This approach is identical to most approaches in cognitive science, which eschew neurophysiological details and focus on attempting to characterize putative structure in behaviour [36]. Put differently, quantum theory represents a hypothesis about the principles that could be reflected in cognition. In pursuing a quantum behavioural model, the hypothesis is that all the basic processes in quantum theory broadly map onto mental processes. This may seem ambitious, but it is no more ambitious than, e.g., the hypothesis that cognition is based on the principles of classical probability theory.

A preliminary issue of some importance is to consider why one might believe that quantum theory has behavioural relevance. That is, in taking a mathematical framework like quantum theory, why do we think that it might offer a plausible psychological model? Currently, in psychology, we do not know. We can simply observe that there appear to be empirical data which are modelled fairly naturally by quantum principles (such as the conjunction fallacy that we focus on just below). Whether the behavioural relevance of quantum theory as a whole can be justified or not is currently unclear. Maybe some, e.g., informational argument will emerge that reveals quantum representations as particularly suitable for mental processes. Or maybe we will have to conclude that there is no behavioural relevance to quantum theory as a whole, rather there are individual heuristics, which sometimes look like quantum theory. Notwithstanding these considerations/concerns, if we accept the relevance of quantum theory in cognition, then there are several novel empirical predictions (see shortly below). The exploration and approximate confirmation of such predictions has provided novel, empirical directions and insights in behavioural sciences.

Even though we will shortly focus on the conjunction fallacy, we note that quantum cognitive models have been pursued across a range of applications. Many of these models have been developed in similar ways, employing, for example, similar assumptions concerning dimensionality, dynamics, the way the initial state is set, etc. A fairly comprehensive overview of quantum cognitive models up to 2011 is provided in [37] and the explanation for the conjunction fallacy below is an example of the kind of models in that overview (the analysis concerning the rational status of the conjunction fallacy is subsequent to this important overview, however). Consider projectors for the bank teller and feminist properties, denoted as P_{BT} and P_F. Assume that the mental state of a participant prior to answering the Linda questions is a pure state denoted by $|\psi\rangle$. Then, we

have that $|P_{BT}|\psi\rangle|^2 = |P_{BT}I|\psi\rangle|^2 = |P_{BT}(P_F + P_{\sim F})|\psi\rangle|^2 = |(P_{BT}P_F + P_{BT}P_{\sim F})|\psi\rangle|^2 = |P_{BT}P_F|\psi\rangle|^2 + |P_{BT}P_{\sim F}|\psi\rangle|^2 + \langle P_{BT}P_{\sim F}\psi|P_{BT}P_F\psi\rangle + \langle P_{BT}P_F\psi|P_{BT}P_{\sim F}\psi\rangle$. This can be written as $\text{Prob}(BT) = \text{Prob}(F \& \text{ then } BT) + \text{Prob}(\sim F \& \text{ then } BT) + \Delta$, an expression which resembles the law of total probability plus Δ. The latter is an interference term which allows violations of the law of total probability in general and, specifically, it can allow $\text{Prob}(BT) < \text{Prob}(F \& \text{ then } BT)$. Note, for incompatible questions we have assumed that conjunctions are meaningful only in a sequential form, that is, instead of $\text{Prob}(F \& BT)$, we have to write $\text{Prob}(F \& \text{ then } BT)$; in the latter term we make an explicit assumption about the specific order in which the questions are assessed.

Why are we restricted to this sequential form of processing for conjunctions of incompatible questions? The principle of 'unicity' is the idea that all events or questions relevant to a situation can be described within a single system of coordinates, that is, with a single framework. Unicity implies that any combinations of questions can be resolved concurrently, so that there is no need to adopt sequential computation. Chichilnisky [38] formulated a topological condition when all events/questions can be expressed within a common framework. Chichilnisky [38–40] derived an impossibility theorem showing that the space of frameworks from a finite Hilbert space (the vector space used in quantum representations) will violate unicity. This then shows, in a formal way, that we cannot represent in a single framework all possible questions, if we commit to a quantum approach. Incompatibility entails a violation of unicity and will force this sequential approach to the computation of, e.g., conjunctions.

These ideas can be illustrated for the case of the conjunction fallacy with a caricature example based on a two-dimensional space, such that all projectors are rays, as in Figure 1. We can further assume a real space (a complex space is not needed even in the general case), so that the dot product between vectors x, y is given by $\langle x|y\rangle = \cos\theta$. Then, $\text{Prob}(BT) = |P_{BT}|\psi\rangle|^2 = |BT\rangle\langle BT|\psi\rangle|^2 = |\langle BT|\psi\rangle|^2 = \cos^2(\pi/2 - s) = \sin^2(s)$. Additionally, $\text{Prob}(F \& \text{ then } BT) = |P_{BT}P_F|\psi\rangle|^2 = |BT\rangle\langle BT|F\rangle\langle F|\psi\rangle|^2 = |\langle BT|F\rangle\langle F|\psi\rangle|^2 = \sin^2(s + t)\cos^2(t)$. A conjunction fallacy exists as long as $\sin^2(s + t)\cos^2(t) > \sin^2(s)$ and there are clearly several angles which satisfy this requirement.
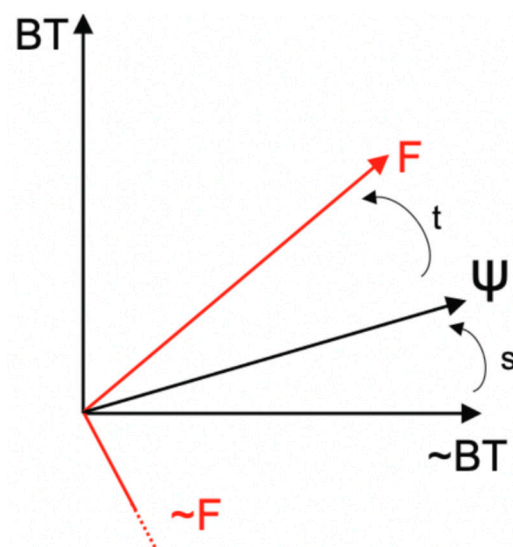
Figure 1. A caricature representation of a quantum model based on one-dimensional projectors, on a two-dimensional space.

Authors in [41] provided a quantum model for the conjunction fallacy, and a range of related findings (e.g., disjunction errors, unpacking effects, etc.) on the basis of these ideas. More generally, quantum theory has been employed in analogous ways in several cases in cognitive science, including in conceptual combination [42], perception [43], memory [44], question order effects [45] and of course decision making [46–50]. Quantum cognitive mod-

els initially focussed on re-explaining puzzling findings, which have resisted compelling classical explanations, like the conjunction fallacy; more recently, the focus has been on the generative potential of quantum models, such as constructive influences [46,49], a quantum Zeno effect [50], the QQ equality [45], etc. (overviews in [37,51,52]).

Regarding the conjunction fallacy, what is exactly the psychological explanation offered? The idea is that from the baseline perspective of the Linda story participants find it very hard to imagine Linda as a bank teller. However, in assessing the conjunction between F and then BT, participants first accept the F property, which is, after all, very likely for Linda. In doing so, they lose some of the initial information about Linda encoded in $|\psi\rangle$ and gain some new insights about her in relation to her characterization as a F. Feminists can have of course many different professions and, from this new feminist perspective it is somewhat easier to accept Linda as a BT. The quantum model for the conjunction fallacy makes two assumptions, that the questions of BT and F are incompatible and that, in the conjunction, the more likely conjunct, F, is evaluated first. Both assumptions are reasonable, but it has to be pointed out that the assumption of incompatibility can be hard to justify independently.

Is the quantum model for the conjunction fallacy a good one? The strength of the model rests primarily in its capacity to cover not just the conjunction fallacy, but a related set of fallacies, including unpacking effects and disjunction errors. With fairly minor modifications it can also cover question order effects and fallacies in relation to probabilistic updating [53]. It is this general scope of explanation which we believe argues in favour of the quantum approach, more so than alternative models. In addition, arguably, the quantum representations and processes could be said to reveal more insight about the problem, than comparable Bayesian approaches, especially if the latter require conditionalizing on context/frame of mind to accommodate results.

Having briefly introduced quantum theory in cognition, a question is why employ the label 'quantum' at all, instead of, e.g., 'projective linear algebra'. The reason is because this seems like the most parsimonious way to summarise the set of assumptions in quantum cognitive models. These assumptions include projective linear algebra, but also collapse, the implications for superpositions, Luder's law, etc. No other connection to quantum mechanics is stated or implied in work with quantum cognitive models.

4. Rationality in the Conjunction Fallacy?

The quantum model for the conjunction fallacy should make it clear that what looks like a conjunction fallacy can be allowed within quantum theory, that is, quantum theory allows situations where $\text{Prob}(A) < \text{Prob}(A \text{ \& then } B)$. Of course, classically, unless we introduce post hoc conditionalizations, it is impossible to have $\text{Prob}(A) < \text{Prob}(A \text{ \& } B)$. Therefore, a person stating that $\text{Prob}(A) < \text{Prob}(A \text{ \& then } B)$ is not necessarily incorrect, as we can conjecture that this person employs quantum, instead of classical, probabilities and representations. Correctness, however, is very different from rationality. Correctness implies consistency with some principle or a set of principles. For example, here we have been exploring consistency with the principles of quantum theory and observed that the conjunction fallacy is correct when considered against these principles. However, another decision maker might adopt another principle, for example, 'to respond yes to all questions'. According to this criterion, very many judgments could be said to be correct, but of course the criterion itself (and so the judgments) can hardly be said to be rational. The point is that correctness is not equivalent to rationality. Arguably, correctness is more narrow, since it depends on a particular criterion (which might be arbitrary, as the example just above shows). By contrast, rationality has to be accompanied by a justification that endows the decision prescription with rational status.

Interestingly, quantum theory can be shown to be consistent with the Dutch Book Theorem in exactly the same way that classical theory is [54]. How can this be possible? The central idea is that we have to create multiple copies of questions, depending on whether they are evaluated in isolation or after other, incompatible questions. This is because

every time a question is resolved (e.g., as would be the case in a sequential conjunction), we essentially change basis set. With each different basis set, we are in a different part of the supporting partial Boolean algebra–essentially, every time a question is resolved it is as if we start operating in a new probability space [55]. Therefore, in the case of the conjunction fallacy, the judgment $\text{Prob}(A) < \text{Prob}(B \,\&\, \text{then } A)$ can be considered $\text{Prob}(A) < \text{Prob}(B \,\&\, A_B)$. The latter inequality is allowed in classical probability theory too and is clearly not problematic. Therefore, why not just employ classical probabilities and simply introduce different copies of the same question whenever we are faced with a result like the conjunction fallacy? Adopting such an approach with classical probability theory would not be much different from post hoc conditionalization. While we will end up with expressions which are not strictly speaking incorrect, the explanatory value of such expressions would be questionable. By contrast, quantum probability theory provides a more constrained and elaborate way of dealing with different probability spaces.

In a more practical sense, the rationality of quantum theory can be expressed as a protection from Dutch books in the more limited setting where we are only prepared to gamble on a subset of possible events. For example, we might be prepared to place bets on which of the set {A & B, A & ~B, ~A & B, ~A & ~B} might occur, but not on the set {A & B, ~A & B, B}. This could be a sensible restriction even classically if we believe that measurement of a system can in general disturb it, so that it matters what happened prior to the measurement of B. In that sense the notion of rationality implied by quantum theory is very natural in any approach where we reject the strong positivist ideal that measuring a system has no influence on it.

If the conjunction fallacy is rational according to one framework for probabilistic inference, but irrational according to another, then the question of rationality boils down to deciding which framework is more applicable. In physics, the use of quantum theory goes in hand with assumptions about the fundamental nature of observables of microscopic particles. Such assumptions cannot really be employed in behavioural sciences. However, if we are to apply quantum theory at all in behavioural science, then we need some way of deciding whether two questions, such as whether Linda is a bank teller and whether she is a feminist, are compatible or incompatible. There is some work indicating that incompatibility can be related to familiarity with the questions and style of cognitive processing [48]. In addition, we think that incompatibility can be understood in terms of whether a question can have a contextual influence on another one or change the state that is relevant for evaluating the other one. For example, consider a question about bugs. Depending on whether this question is preceded by a question about holidays vs. spies, the bugs question will acquire different meanings. As another example, suppose that baby Eve's parents are wondering whether she sleeps on her back at 10:00 p.m. and 11:00 p.m. Baby Eve in general likes sleeping on her back, but when her parents check on her that disturbs her and she rolls over. One can easily see how apparent conjunction fallacies can arise when considering corresponding questions.

In general, we think that incompatibility has meaning for questions in our macroscopic environment exactly in such ways, either because of contextuality or because of prior questions disturbing the 'system' for subsequent ones. Then, in a case like the Linda conjunction fallacy, because it is difficult to motivate incompatibility either as contextuality or from disturbing questions, we have to conclude that the judgment is still irrational–though it is possible that the irrationality arises because questions which should be treated as compatible are treated as incompatible, as opposed to as an error in probabilistic judgment [54].

A final consideration concerns the present approach to rationality itself. In this work we have focussed on classical probability theory, as this is the currently predominant view of rationality in behavioural sciences [1–3]. However, what about alternative approaches, such as classical logic? Briefly, the effort to understand rational thought using classical logic came under scrutiny with Wason's experimental work, which revealed sharp discrepancies between human behaviour and logical prescription in a simple reasoning task [56,57]. More sophisticated approaches to logic may well reveal a promising candidate for rationality, but

this has yet to happen. As an aside, work in quantum logic, such as the History Projector Operator (HPO) formalism, offers promise regarding behavioural theory, but such ideas have yet to find their way to cognitive models.

5. Concluding Comments

The question of rationality has been one of the most significant ones in scientific inquiry, both because of its vast importance (e.g., in relation to political, legal, or medical decision making, all cases where it is often critically important to reach decisions as rational as possible) but also because of its centrality for the way we understand the potential of human intellect. We think that quantum probability theory allows a major step in the way we can comprehend rationality. Briefly, the heart of rationality is still classical probability theory. However, we can no longer assume an all-encompassing set of questions which are all compatible. Instead, we have to adopt a fragmented view, such that there are sets of questions; compatibility exists only within sets and between sets we have incompatibility. In fact, there has been independent support for the idea that our knowledge is 'partitioned' in a way that across partitions we appear to lose our capacity for consistent inference [58].

A key question is why we might employ incompatible representations when the questions are neither contextual nor have capacity to disturb each other, as it appears to be the case in the Linda problem. The observation that lack of familiarity appears to make incompatible representations more likely [48] suggests that it takes 'effort' to build compatible representations. Indeed, it is straightforward to show that the dimensionality of a probability space required to represent the same set of questions as fully compatible vs. fully incompatible can vary dramatically. Therefore, is the apparent use of incompatible representations in cognitive process a matter of simplification, e.g., when there is less need to invest the effort to build compatible representations? We think this is a highly pertinent question for future work [59].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Griffiths, T.L.; Chater, N.; Kemp, C.; Perfors, A.; Tenenbaum, J.B. Probabilistic models of cognition: Exploring representations and inductive biases. *Trends Cogn. Sci.* **2010**, *14*, 357–364. [CrossRef]
2. Oaksford, M.; Chater, N. *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*; Oxford University Press: New York, NY, USA, 2007.
3. Tenenbaum, J.B.; Kemp, C.; Griffiths, T.L.; Goodman, N.D. How to Grow a Mind: Statistics, Structure, and Abstraction. *Science* **2011**, *331*, 1279–1285. [CrossRef] [PubMed]
4. De Finetti, B.; Machi, A.; Smith, A. *Theory of Probability: A Critical Introductory Treatment*; Wiley: Hoboken, NJ, USA, 1993.
5. Azevedo, E.M.; Gottlieb, D. Risk-neutral firms can extract unbounded profits from consumers with prospect theory preferences. *J. Econ. Theory* **2012**, *147*, 1291–1299. [CrossRef]
6. Wakker, P.P. *Prospect Theory for Risk and Ambiguity*; Cambridge University Press: New York, NY, USA, 2010.
7. Perfors, A.; Tenenbaum, J.B.; Griffiths, T.L.; Xu, F. A tutorial introduction to Bayesian models of cognitive development. *Cognition* **2011**, *120*, 302–321. [CrossRef] [PubMed]
8. Gilboa, I. *Theory of Decision under Uncertainty*; Cambridge University Press: New York, NY, USA, 2009.
9. Moore, D.W. Measuring New Types of Question-Order Effects. *Public Opin. Q.* **2002**, *66*, 80–91. [CrossRef]
10. Kahneman, D. *Thinking Fast and Slow*; Penguin: London, UK, 2012.
11. Sloman, S.A. The empirical case for two systems of reasoning. *Psychol. Bull.* **1996**, *119*, 3–22. [CrossRef]
12. Bergus, G.R.; Chapman, G.B.; Levy, B.T.; Ely, J.W.; Oppliger, R.A. Clinical Diagnosis and the Order of Information. *Med. Decis. Mak.* **1998**, *18*, 412–417. [CrossRef] [PubMed]

13. McKenzie, C.R.M.; Lee, S.M.; Chen, K.K. When negative evidence increases confidence: Change in belief after hearing two sides of a dispute. *J. Behav. Decis. Mak.* **2002**, *15*, 1–18. [CrossRef]

14. Pennington, N.; Hastie, R. Evidence evaluation in complex decision making. *J. Personal. Soc. Psychol.* **1986**, *51*, 242–258. [CrossRef]

15. Trueblood, J.S.; Busemeyer, J.R. A Quantum Probability Account of Order Effects in Inference. *Cogn. Sci.* **2011**, *35*, 1518–1552. [CrossRef] [PubMed]

16. Shafir, E.; Tversky, A. Thinking through uncertainty: Nonconsequential reasoning and choice. *Cogn. Psychol.* **1992**, *24*, 449–474. [CrossRef]

17. Pothos, E.M.; Perry, G.; Corr, P.J.; Matthew, M.R.; Busemeyer, J.R. Understanding cooperation in the Prisoner's Dilemma game. *Pers. Individ. Differ.* **2011**, *51*, 210–215. [CrossRef]

18. Broekaert, J.; Busemeyer, J.; Pothos, E. The Disjunction Effect in two-stage simulated gambles. An experimental study and comparison of a heuristic logistic, Markov and quantum-like model. *Cogn. Psychol.* **2020**, *117*, 101262. [CrossRef]

19. Townsend, J.T.; Silva, K.M.; Spencer-Smith, J.; Dror, I.E.; Stevenage, S.V.; Wenger, M.J. Exploring the relations between categorization and decision making with regard to realistic face stimuli. *Pragmat. Cogn.* **2000**, *8*, 83–105. [CrossRef]

20. Tversky, A.; Kahneman, D. Extensional versus intuitive reasoning: The conjuctive fallacy in probability judgment. *Psychol. Rev.* **1983**, *90*, 293–315. [CrossRef]

21. Tentori, K.; Bonini, N.; Osherson, D. The conjunction fallacy: A misunderstanding about conjunction? *Cogn. Sci.* **2004**, *28*, 467–477. [CrossRef]

22. Schwarz, N. Attitude Construction: Evaluation in Context. *Soc. Cogn.* **2007**, *25*, 638–656. [CrossRef]

23. Grice, H.P. Logic and conversation. In *Syntax and Semantics 3: Speech Acts*; Cole, P., Morgan, J.L., Eds.; Academic Press: New York, NY, USA, 1975; pp. 41–58.

24. Dulany, D.E.; Hilton, D.J. Conversational Implicature, Conscious Representation, and the Conjunction Fallacy. *Soc. Cogn.* **1991**, *9*, 85–110. [CrossRef]

25. Wedell, D.H.; Moro, R. Testing boundary conditions for the conjunction fallacy: Effects of response mode, conceptual focus, and problem type. *Cognition* **2008**, *107*, 105–136. [CrossRef]

26. Moro, R. On the nature of the conjunction fallacy. *Synthese* **2009**, *171*, 1–24. [CrossRef]

27. Tentori, K.; Crupi, V.; Russo, S. On the determinants of the conjunction fallacy: Probability versus inductive confirmation. *J. Exp. Psychol. Gen.* **2013**, *142*, 235–255. [CrossRef]

28. Busemeyer, J.R.; Wang, J.; Pothos, E.M.; Trueblood, J.S. The conjunction fallacy, confirmation, and quantum theory: Comment on Tentori, Crupi, and Russo (2013). *J. Exp. Psychol. Gen.* **2015**, *144*, 236–243. [CrossRef]

29. Costello, F.; Watts, P. Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychol. Rev.* **2014**, *121*, 463–480. [CrossRef]

30. Zhu, J.; Sanborn, A.N.; Chater, N. The Bayesian sampler: Generic Bayesian inference causes incoherence in human probability judgments. *Psychol. Rev.* **2020**, *127*, 719–748. [CrossRef]

31. Sorkin, R.D. Quantum mechanics as quantum measure theory. *Mod. Phys. Lett. A* **1994**, *9*, 3119–3127. [CrossRef]

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

33. Gigerenzer, G.; Goldstein, D.G. Reasoning the fast and frugal way: Models of bounded rationality. *Psychol. Rev.* **1996**, *103*, 650–669. [CrossRef] [PubMed]

34. Lieder, F.; Griffiths, T.L. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behav. Brain Sci.* **2019**, *43*, 1–85. [CrossRef] [PubMed]

35. Litt, A.; Eliasmith, C.; Kroon, F.W.; Weinstein, S.; Thagard, P. Is the Brain a Quantum Computer? *Cogn. Sci.* **2006**, *30*, 593–603. [CrossRef]

36. Marr, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*; W. H. Freeman: New York, NY, USA, 1982.

37. Busemeyer, J.R.; Bruza, P. *Quantum Models of Cognition and Decision Making*; Cambridge University Press: New York, NY, USA, 2011.

38. Chichilnisky, G. *The Topology of Quantum Theory and Social Choice*; Working Paper; Columbia University: New York, NY, USA, 2016.

39. Chichilnisky, G. *A Topological Characterization of the Space of Events and Frameworks of R^n*; Working Paper; Stanford University: Stanford, CA, USA, 2015.

40. Chichilnisky, G. Quantum Theory and Social Choice. In Proceedings of the XXVII European Workshop on General Equilibrium Theory, Paris, France, 28 June 2018.

41. Busemeyer, J.R.; Pothos, E.M.; Franco, R.; Trueblood, J.S. A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* **2011**, *118*, 193–218. [CrossRef]

42. Aerts, D.; Aerts, S. Applications of quantum statistics in psychological studies of decision processes. *Found. Sci.* **1995**, *1*, 85–97. [CrossRef]

43. Atmanspacher, H.; Filk, T. A proposed test of temporal nonlocality in bistable perception. *J. Math. Psychol.* **2010**, *54*, 314–321. [CrossRef]

44. Bruza, P.; Kitto, K.; Nelson, D.; McEvoy, C. Is there something quantum-like about the human mental lexicon? *J. Math. Psychol.* **2009**, *53*, 362–377. [CrossRef] [PubMed]

45. Wang, Z.; Solloway, T.; Shiffrin, R.M.; Busemeyer, J.R. Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9431–9436. [CrossRef] [PubMed]

46. Kvam, P.D.; Pleskac, T.J.; Yu, S.; Busemeyer, J.R. Interference effects of choice on confidence: Quantum characteristics of evidence accumulation. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10645–10650. [CrossRef] [PubMed]

47. Pothos, E.M.; Busemeyer, J.R. A quantum probability explanation for violations of 'rational' decision theory. *Proc. R. Soc. B* **2009**, *276*, 2171–2178. [CrossRef]

48. Trueblood, J.S.; Yearsley, J.M.; Pothos, E.M. A quantum probability framework for human probabilistic inference. *J. Exp. Psychol. Gen.* **2017**, *146*, 1307–1341. [CrossRef] [PubMed]

49. White, L.C.; Pothos, E.M.; Jarrett, M. The cost of asking: How evaluations bias subsequent judgments. *Decision* **2020**, *7*, 259–286. [CrossRef]

50. Yearsley, J.M.; Pothos, E.M. Zeno's paradox in decision-making. *Proc. R. Soc. B* **2016**, *283*, 20160291. [CrossRef]

51. Haven, E.; Khrennikov, A. *Quantum Social Science*; Cambridge University Press: New York, NY, USA, 2013.

52. Pothos, E.M.; Busemeyer, J.R. Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* **2013**, *36*, 255–274. [CrossRef] [PubMed]

53. Basieva, I.; Pothos, E.; Trueblood, J.; Khrennikov, A.; Busemeyer, J. Quantum probability updating from zero priors (by-passing Cromwell's rule). *J. Math. Psychol.* **2017**, *77*, 58–69. [CrossRef]

54. Pothos, E.M.; Busemeyer, J.R.; Shiffrin, R.M.; Yearsley, J.M. The rational status of quantum cognition. *J. Exp. Psychol. Gen.* **2017**, *146*, 968–987. [CrossRef] [PubMed]

55. Hughes, R.I.G. *The Structure and Interpretation of Quantum Mechanics*; Harvard University Press: Cambridge, MA, USA, 1989.

56. Wason, P.C. Reasoning about a Rule. *Q. J. Exp. Psychol.* **1968**, *20*, 273–281. [CrossRef] [PubMed]

57. Wason, P.C.; Johnson-Laird, P.N. *The Psychology of Reasoning: Structure and Content*; Harvard University Press: Cambridge, MA, USA, 1972.

58. Lewandowsky, S.; Roberts, L.; Yang, L.-X. Knowledge partitioning in categorization: Boundary conditions. *Mem. Cogn.* **2006**, *34*, 1676–1688. [CrossRef] [PubMed]

59. Pothos, E.M.; Lewandowsky, S.; Basieva, I.; Barque-Duran, A.; Tapper, K.; Khrennikov, A. Information overload for (bounded) rational agents. *Proc. R. Soc. B* **2021**, *288*, 20202957. [CrossRef]