Quantum Cognition: A New Theoretical Approach To Psychology

Peter D. Bruza

Zheng Wang

Jerome R. Busemeyer

Peter D. Bruza, Information Systems School, GPO Box 2434, Queensland University of Technology, Brisbane 4001, Australia

Zheng Wang, School of Communication, 3145 Derby Hall, 154 N. Oval Mall, The Ohio State University, Columbus OH, 43210, USA

Jerome R. Busemeyer, Department of Psychological and Brain Sciences, 1101 E. 10th St., Indiana University, Bloomington IN, 48705, USA

Corresponding author: Bruza, P.D. (p.bruza@qut.edu.au)

Keywords: quantum probability, Bayesian probability, complementarity, incompatibility, superposition, human judgment

Abstract

What kind of probability theory best describes the way humans make judgments under uncertainty and decisions under conflict? Although rational models of cognition have become prominent and achieved much success, they adhere to the laws of classical probability theory despite the fact that human reasoning does not always conform to these laws. For this reason, we have seen the recent emergence of models based on an alternative probabilistic framework drawn from quantum theory. These quantum models show promise in addressing cognitive phenomena that have proven recalcitrant to modeling by means of classical probability theory. This review compares and contrasts probabilistic models based on Bayesian or classical vs. quantum principles, and highlights the advantages and disadvantages of each approach.

A new approach to cognitive science

Cognitive scientists have long struggled to form a comprehensive understanding of how humans make judgments and decisions under conflict and uncertainty. Decades of research have seen two approaches crystalize to the surface—“heuristic” and “rational.” The heuristic approach is firmly rooted in Herbert Simon’s notion of bounded rationality [1]. This approach posits that, to make judgments and decisions, people tend to employ simple heuristics (e.g., representativeness, anchoring-and-adjustment, take-the-best), which may not always seem rational. This can be viewed as a “bottom up” inductive process in the sense that humans learn simple *ad hoc* rules that can be effective or not depending on the environmental conditions [2-4]. In stark contrast, the rational approach is founded on theories of subjective probability and expected utility [5]. This approach posits that people can derive inferences from the Bayes rule and decisions from the expected utility rule in a rational manner. This can be viewed as a “top down" deductive process, wherein the same basic axioms can be used to derive inferences and utilities across all environmental conditions [6, 7]. Recently, a third approach, called "quantum cognition," has emerged [8-10]. In common with the heuristic approach, it assumes that the human decision maker is subject to bounded rationality. Moreover, like the rational approach, inferences used for decisions are derived from basic axioms that define a probability theory. However, the axioms are different from those employed by the Bayesian approach, and consequently, so are the inferences that are derived from it.

To what probability theory does quantum cognition subscribe? This question may come as a surprise to some readers, since by and large, cognitive scientists have been exposed to a single probability theory—what we will call classical (more technically, Kolmogorov) probability theory, on which Bayesian models rest. There are, however, several viable probability theories upon which to build probabilistic models of cognition [11]. Quantum cognition is one such alternative that is increasingly gaining attention [8-10]. What makes quantum cognition controversial is that its associated probability theory was developed within the field of quantum physics. This immediately raises skepticism: How could a probability theory governing the behavior of subatomic phenomena have anything meaningful to say about cognitive phenomena, such as decision making? Indeed, why even consider an alternative to classical probability theory? How do the classical and quantum probability theories differ, what critical assumptions do they make, and how can we decide which theory is superior for modeling cognitive phenomena? This review will be guided by such questions. First, we present some compelling psychological reasons why certain abstract principles in quantum theory are relevant to cognitive phenomena. With this foundation in place, we will then briefly but systematically compare the assumptions underlying classical and quantum probability theories. This comparison provides the intuition regarding how each theory makes different predictions, and how their prescriptions of what is rational differ. Finally, the conclusion summarizes the critical differences between quantum and classical models, and identifies the reasons why they differ in this way.

From quantum physics to quantum cognition

Less than 20 years ago, a group of scientists pioneered the bold idea of applying the abstract principles from quantum theory outside of physics into the field of human cognition [12-15]. They did not proceed from the assumption that there is something quantum-like going on in the brain, but instead drew inspiration from both the mathematical structure of quantum theory and its associated dynamic principles.

It was the mathematician John von Neumann [16] who axiomatized quantum theory in the 1930's. Regarding these axioms, he stated, "The set theoretical situation of logics is replaced by the machinery of projective geometry, which in itself is quite simple." ([17], p.244) He was referring to the fact that the logic underlying classical probability theory is prescribed by Boolean algebra, because of its set-theoretic foundation. The associated Boolean logic implies that events can always be combined (e.g., via logical conjunction). The consequence, however, is that logical conjunction is commutative, so expressing the combined event "*A and B*" is the same as stating "*B and A*." In other words, the order of the events does not matter. In contrast, inherent in von Neumann's "machinery of projective geometry" is non-commutativity: The sequence of events "*A and B*" is not necessarily the same as the sequence of "*B and A*."

To illustrate the idea of commutativity, consider an example of public opinion polls. Polling companies know well that asking respondents to judge the trustworthiness of a politician A followed by another politician B can generate different results compared to when the politicians are judged in the reverse order [18-19]. This example highlights the fact that human judgment is not necessarily commutative. To account for this non-commutativity, a classical model needs to include the sequence of evaluations as part of the description of the event (e.g., "*A* and *B" and* "*A* before *B"*). This point will be taken up in more detail later. Meanwhile, we can begin to appreciate why the mathematical structure of quantum theory might lend itself to the modeling of human cognition as many cognitive phenomena crucially depend on the sequence or order of the cognitive processes and measurements (e.g., judgments and decisions), and quantum theory was initially developed to address order effects of measurements in physics. Of course, not all judgments and decisions are sensitive to order. For example, if one decision is whether to drive a car to work or take a bus, and the other is whether to have a dog or cat as a family pet, then we typically would not expect the outcomes of these decisions to be sensitive to the sequence in which the decisions are taken. In principle, these decisions can be evaluated simultaneously and evaluating one does not interfere with what we think about the other, so the same outcomes are produced regardless of the order of decisions.

The preceding two examples provide a simple introduction to the principle of complementarity ([20], see Glossary). When this principle is transmigrated into quantum models of cognition, it states that we need to distinguish between two kinds of events, compatible and incompatible, in order to understand how humans reason under uncertainty. More specifically, two events (e.g., two questions), *A and B*, are deemed "compatible" if they can be considered simultaneously. That is, the order in which they are considered does not matter to the finally observed or measured values. Conversely, events *A and B* are defined as "incompatible" if they cannot be considered simultaneously but thus have to be taken sequentially. In this case the order does matter to the finally observed or measured values. The complementarity principle essentially posits the existence, or even the prevalence, of incompatible measures. As a note in passing, Niels Bohr, one of the founding fathers of quantum theory, introduced the principle of complementarity into quantum physics. He originally pointed out that this principle is what distinguishes quantum from classical probability theory [21]. This distinction will be made more explicit in the following section.

The principle of complementarity leads directly to the most well known principle of quantum theory—the uncertainty principle [21]. This principle holds that when we are certain about a quantum particle's position, we are necessarily uncertain about its momentum, and *vice versa*. Placed in a psychological context, the uncertainty principle becomes relevant because a person's understanding of two events, such as two different politicians or two different perspectives on a matter, requires changing from one point of view to another, and the two points of view can imply incompatibility. For example, suppose you are planning a trip with a friend and need to select a trip from a set of alternatives. For a given trip, you can judge how much you will enjoy the trip yourself from your own perspective, and also judge how much your friend will enjoy it from his or her perspective. The intuition behind the uncertainty principle is that after clarifying your friend's position, you may become uncertain about your own; likewise, after clarifying your own position, you may become uncertain about your friend's. In other words, the uncertainty principle entails that it is not possible to be simultaneously decided on the matter across both your and your friend's perspectives. Such events (e.g., perspectives) are deemed incompatible simply because they cannot be jointly considered simultaneously. The reason for this relates to another basic principle of quantum theory, namely superposition. In the context of our example, superposition means that when you are decided about a matter from one perspective (e.g., you judge that you will enjoy the trip), your cognitive state has to be dispersed, or indefinite, with respect to the other perspective (i.e., your friend's enjoyment of the trip).

The principle of complementarity can be viewed as a constraint that bounds rationality. This principle, along with the uncertainty and superposition principles, lead to key differences between classical and quantum theories of probability [22]. These principles, we will argue, are critical for understanding the limitations on human reasoning and decision making.

Comparison of classical and quantum probability theories

Classical probability theory developed over centuries from a grounding based on classical physics. However, an axiomatic foundation for classical probability theory was not formulated until the 20th century by Andrey Kolmogorov [23]. Modern applications of Bayesian inference in cognitive science are derived from Kolmogorov's theory. As we mentioned earlier, the axiomatic foundation for quantum theory was formulated by John von Neumann around the same time—an interesting coincidence [16]. Once the abstract mathematical principles were identified and separated out from the physical content, it became clear that what quantum physicists had discovered was a new probability theory [24]. In this section, we will compare the two probability theories, side by side, to gain an understanding of where they are similar and where they differ. This understanding is important when one considers the application of these theories for modeling cognitive phenomena.

Classical probability theory

Classical probability theory is based on events, which in Kolmogorov's theory, correspond to subsets of a universal set. By way of illustration, let *A* denote the event that you think you would enjoy the trip being considered, and let *B* denote the event that you think your friend would enjoy the same trip. According to Kolmogorov's probability theory, *A and B* are subsets of a universal set that includes, as elements, all combinations of answers regarding what you think about your and your friend's enjoyment of the trip. Previously we saw how events can be combined via conjunction (i.e., *A and B*). In Kolmogorov's theory, this conjunction is represented as set intersection, $\boldsymbol{A} \cap \boldsymbol{B}$, which in our example represents the conjunction that you think both you and your friend would enjoy the trip. In addition, events may be combined by set union, $\boldsymbol{A} \cup \boldsymbol{B}$, which in our example represents the disjunction that you think you *or* you think your friend would enjoy the trip. These definitions imply a Boolean logic for events. Given that *A* is an event and *B* is an event, then $\boldsymbol{A} \cap \boldsymbol{B}$ is an event, and so is $\boldsymbol{A} \cup \boldsymbol{B}$. Recall that this logic of events is commutative, that is, $\boldsymbol{A} \cap \boldsymbol{B} = \boldsymbol{B} \cap \boldsymbol{A}$. Another important property of this logic is the distributive axiom: $\boldsymbol{A} = \boldsymbol{A} \cap (\boldsymbol{B} \cup \sim\boldsymbol{B}) = (\boldsymbol{A} \cap \boldsymbol{B}) \cup (\boldsymbol{A} \cap \sim\boldsymbol{B})$, where ~*B* denotes the negation of event *B*. In addition, classical models of cognition generally assume that a person has a personal probability function *p*, which is used to assign probabilities to events according to the rules summarized in Box 1.

Quantum probability theory

The corresponding foundations in quantum theory are markedly different. John von Neumann replaced the set theoretic structure of classical theory with the projective geometric structure of vector spaces using two separate steps [16]. First, the definition of events was changed from one based on subsets of a universal set to another based on subspaces of a vector space. Second, instead of using a function *p* to assign probabilities to events, von Neumann used a state vector, *S*, to assign these probabilities as described in Box 2.

Mathematically, a vector is an arrow lying within a vector space. The state vector, *S*, might seem like an abstract idea, but when applied to cognitive science, it is analogous to the distributed input across nodes in a connectionist neural network [25-26]. Mathematically, a subspace is a geometric object (e.g., a ray, a plane, a hyperplane) lying within a vector space, and each subspace corresponds to a projector that maps the state vector *S* onto the subspace. A projector might seem like another abstract idea, but it also has a natural interpretation in cognitive science: It maps the distributed input into a distributed output analogous to a connectionist neural network [25-26] (see Figure 1). This mapping process is called projecting, and the resulting output is called the projection. The probability assigned to an event equals the squared length of the projection.

Let us provide some accompanying intuition by referring to Figure 2. Suppose the state vector *S* represents your beliefs about enjoyment of a trip. First, consider the question about trip enjoyment from your self-perspective, as shown in Panel a. The probability of event *A* (you think you will enjoy the trip) *or* event *B* (you think you will be neutral about the trip) is determined by the projection of the state vector *S* onto the plane spanned by the *A*, *B* axes, represented by the blue vector *T*. Projection is like shining a light from above and seeing the length of the shadow onto the plane (i.e., projection vector *T* is the shadow in Figure 2-a); the longer the shadow the higher the probability. Formally, the squared length of the projection *T* constitutes the probability of deciding to choose either event *A* or *B* (which equals .833 in the example). Similarly, the probability of event *C* (you think you will not enjoy the trip) is determined by the squared length of the projection of *S* onto the axis *C* (which equals .167 in the example).

Now consider the trip's enjoyment from the perspective of your friend by referring to Panel b of Figure 2. This change in perspective is accomplished by rotating the *A*, *B*, and *C* axes from the self-perspective (Panel a) to form the three axes *U* (you judge your friend will enjoy the trip), *V* (you judge your friend will be neutral about the trip), and *W* (you judge your friend will not enjoy the trip) from the other-perspective (Panel b). The degree of rotation between the two perspectives is determined by the similarity between your self-perspective and friend-perspective (i.e., less rotation means the perspectives are more similar). The probabilities of answers from your friend's perspective are then determined by projecting the state *S* onto events defined by the *U*, *V*, and *W* axes. For example, the probability that you think your friend will be neutral about the trip equals the squared length of the projection of the state vector *S* onto ray *V* (which equals .812 in the example). Note that if you decide that your friend will be neutral about the trip, then your state vector is updated to align with ray *V*. Observe carefully the consequence of this: You must now necessarily become undecided about your own enjoyment of the trip as rays *A, B* and *C,* that represent your perspective, are all oblique with respect to ray *V*. Technically speaking, being certain about event *V* forces you to be superposed with respect to the events *A*, *B*, and *C*, as determined by the quantum principles of superposition and uncertainty described earlier.

Earlier we introduced Niels Bohr's principle of complementarity, and Figure 2 provides an example of how this principle can be understood in terms of projections. If first we project the state *S* onto the *A*, *B* plane to produce the projection *T*, and then subsequently project *T* onto ray *V*, then the squared length of the final projection equals .375. If instead, we first project the state *S* onto ray *V* and then subsequently project the resulting projection onto the *A*, *B* plane, then we acquire a different final projection with squared length of .406. In other words, reversing the order of projections alters the final probability of two decisions taken in sequence. This in turn implies that the events corresponding to these decisions are incompatible, and consequently these events do not commute.

Note this does not mean that a person mentally rotates some kind of axes in *N*-dimensional spaces. On the contrary, we think this is accomplished implicitly through some interconnected neural network system [25-26], and quantum theory provides algorithmic level predictions for this more complex neural implementation. On the one hand, neural networks are not required to satisfy the axioms of quantum theory; on the other hand, the quantum algorithm could be generated from a neural network [26].

In general, quantum theory defines events as subspaces. Each subspace has a corresponding projector, where the associated logic of events is not necessarily Boolean [27]. In particular, if two events *A*, *B* are incompatible, then the conjunction of events: *A and B*, cannot be defined as they do not commute. Instead, only the sequence of events, *A and then B,* can meaningfully be defined. This constitutes a major difference from classical probability theory where the intersection of events is always defined and events are always commute, even if the events are distinguished by time (e.g., "*A* at time 1" *and* "*B* at time 2" is equivalent to "*B* at time 2" *and* "*A* at time 1"). Further, in quantum theory, if event *A* is incompatible with event *B*, then the distributive axiom also fails because the probability of event *B* is no longer equal to the probability of *A and then B* plus the probability of *~A and then B*.

Empirical support for quantum cognition

Why use quantum probability theory? After all, the prevalence of the Bayesian models is testament to the success of classical probability theory in modeling cognition. Despite this success, however, there has been a steady accumulation of puzzling, even paradoxical, cognitive phenomena that violate the axioms upon which classical probability theory (and hence Bayesian inference) is based. Thus far, these violations have been explained using heuristic rules such as the representativeness heuristic and the anchoring-and-adjustment heuristic. Rather than resorting to heuristics, quantum cognition successfully accounts for these violations using a coherent, common set of principles. Although this review focuses on judgment and decision making, we briefly highlight the expressive power of quantum models by pointing out that they have been applied to a broad range of cognitive phenomena, including perception [28-29], memory [30-32], conceptual combinations [33-36], attitudes [37-38], probability judgments [39-

42], causal reasoning [43], decision making [44-50], and strategic games [51-52]. It is not possible to survey the myriad of applications in this review. We will restrict our attention to a representative set of examples, which intuitively illustrate the basic quantum principles introduced previously.

Probability judgment errors

Decades of research on human judgment have uncovered a full spectrum of so-called irrational phenomena that deviate substantially from what would be considered normatively correct according to classical logic and probability theory [39]. A popular example is "the Linda problem." As describe in Box 3, participants read a description of a liberal female college student named Linda and are asked to judge the probability of a series of statements about her after she graduated. This extensively studied experimental paradigm has shown that human participants consistently rate the probability of the conjunction (being a feminist and bank teller) to be greater than that of its constituent event (being a bank teller). This probability judgment error is so robust that it has been named the conjunction fallacy, and was originally explained by the representativeness heuristic [53].

According to quantum cognition, a key to explaining the conjunction fallacy is the incompatibility between events—in the case of the Linda problem, that Linda is a bank teller and that she is a feminist [39]. Referring to Panel a in Figure 3, suppose the person considers the event that Linda is a bank teller. The probability of this decision is defined by the green projection from the state *S* down onto the red *B* axis (*B* stands for bank teller). Observe that the projection is short, indicating a low probability (.024 in the example). If the person makes this decision, the cognitive state *S* collapses onto the axis *B*, reflecting that they have decided that Linda is a bank teller. Note carefully that as illustrated in the figure, the person must now necessarily be uncertain about whether Linda is a feminist because the axis *B* that corresponds to the current cognitive state is suspended between the two vectors, *F* and $\neg F$ (*F* stands for feminist). The hallmark of incompatibility is the state of indecision from one perspective (e.g., the feminist perspective) when a decision is taken from another (e.g., the bank teller perspective).

Now consider the probability of the conjunction of events that Linda is a feminist and bank teller (*F and B*), which as explained in the previous section, must be considered sequentially (*F and then B*) because *F* and *B* are incompatible. More specifically, the probability of the conjunction is determined by first projecting the state *S* onto the axis *F* and then projecting this result onto the axis *B* (following the black lines in the figure). Inspection of the figure shows that the result of this two-step sequence of projections is longer (a .093 probability in the example) than that obtained from event *B* alone. Therefore, we can intuitively appreciate how these incompatible perspectives naturally allow the conjunction of events to have a higher probability than a single constituent event, thus explaining the conjunction fallacy in a simple and axiomatic way, without recourse to *ad hoc* heuristics.

The basic quantum model underpinning the conjunction fallacy also readily explains the disjunction fallacy [39, 54]. Furthermore, this model makes new *a priori* predictions [39, 41]. Foremost among them is the consequence that incompatible judgments and decisions must entail order effects. More specifically, the conjunction fallacy only occurs when the decision of Linda being a feminist, or not, is evaluated before the decision of whether she is a bank teller. Indeed, the conjunction fallacy has been found to depend on this order [55]. Order effects are central to quantum theory, and we will review them next.

Order effects on attitudes, inference, and causal reasoning

Order effects are pervasive in human cognition, such as in attitude judgments [18-19, 37], inferences [40], and causal reasoning [43]. An interesting example is the results of a 1997 poll in the U.S. concerning public opinions on the characters of Bill Clinton and Al Gore [56]. Half of the 1,002 nationally sampled respondents were asked: "Do you generally think Bill Clinton is honest and trustworthy?" Subsequently, they were asked the same question about Al Gore. The other half of respondents answered exactly the same questions, but in the opposite order. The results exhibited a striking order effect: The proportion saying "yes" to both questions significantly increased when Gore was judged first.

Quantum theory provides a natural account for question order effects in terms of incompatibility [37, 57]. Referring to Panel b in Figure 3, first, consider the probability of "yes" to both questions when Clinton is judged first. This is obtained by first projecting (following the black lines) the state *S* down onto the blue horizontal Clinton axis and then projecting the result to the red Gore axis, which produces a low probability (.173 in the example). Next, consider the probability of "yes" to both questions when Gore is judged first. This is obtained by first projecting (following the green lines) the state *S* to the red Gore axis and then projecting the result down to the blue Clinton axis, which produces a higher probability (.487 in the example).

On the basis of the preceding examples, criticism might arise that the quantum model lacks predictive power and only explains order effects in a *post hoc* manner. This criticism is countered by the fact that the quantum model makes a precise, *a priori* prediction of a very special pattern of order effects, called the QQ equality [37, 57]:

$$[\,p(C\ and\ then\ {\sim}G) - p({\sim}G\ and\ then\ C)] = -[p({\sim}C\ and\ then\ G) - p(G\ and\ then\ {\sim}C)],$$

where for example, $p(C\ and\ then\ {\sim}G)$ is the probability of saying "yes" to Clinton and then "no" to Gore. Basically, the QQ equality predicts a surprising regularity in question order effects.

This is a completely general prediction: It must hold for any pair of questions measured back to back; it must hold for any dimension *N* for the vector space; it must hold for any rotation of the axes; it must hold for compatible and for incompatible events; it must hold regardless of initial state; and it must even hold when averaging across people with different state vectors (see [37, 57] for proof). In short, the QQ equality is an exceptionally strong *a priori* prediction. Recent research has shown that it was statistically supported across 70 national field experiments that examined question order effects using a wide range of topics [37] (see Figure 4).

The quantum model is not the only model that can account for order effects. Another traditional model is the anchoring-and-adjustment model based on the heuristics approach [58]. However, anchoring-adjustment models failed to account for the empirical patterns of order effects and QQ equality observed in the 70 experiments mentioned above [37]. In addition, the anchoring-and-adjustment model fell short of the quantum model when the two were quantitatively compared in experiments investigating order effects on inference [40]. The general finding of the experiments is that recent evidence has greater impact than early evidence [40], which is uncorrelated with memory recall [59]. The anchoring-and-adjustment model failed to even reproduce the correct ordering of the size of the order effects across the experimental conditions, whereas the quantum model accurately reproduced all of the order effects [40]. Similar findings favoring the quantum model were also obtained using a causal reasoning paradigm [43]. Although we have emphasized applications to order effects, quantum theory has important applications to many other cognitive phenomena, including empirical violations of the classical law of total probability (see Box 1), as discussed next.

Violations of rational decision making

One of the most important rational axioms of decision-making in classical theory is called the "sure thing" principle [5]. It states that if you prefer action A over B under the state of the world X, and you also prefer action A over B under the opposite state of the world ~X, then you should prefer action A over B even if the state of the world is unknown. Violations of the sure thing principle produce violations of the basic classical law of total probability, which is the foundation of Bayesian and Markov models, and thus holds an important role in modeling cognition. This principle has been directly tested by making an innovative modification of the well-known prisoners' dilemma (PD) game [60]. In the typical PD game (see Box 4), each player chooses to either cooperate or defect, and the two players move simultaneously without knowing each other's move. The payoffs are arranged so that defection is the dominant choice. The key innovation for testing the sure thing principle was to include trials in the game that informed a player of the opponent's move before acting. What happened was quite surprising: When informed that the opponent defected, the players defected on 97% of the trials; and when informed that the opponent cooperated, the players continued to defect on 84% of the trials; but when the opponent's action was unknown, the players only defected on 66% of the trials [60]. In other words, many players defected when the opponent's action was known, but they changed to cooperate when it was unknown. These results are shown to be robust, and present a clear violation of the law of total probability. According to this law,

$$p(Player\ Defects) = p(Opp\ Defects) \cdot p(Player\ Defects|Opp\ Defects)$$
$$+p(Opp\ Cooperates) \cdot p(Player\ Defects|Opp\ Cooperates).$$

That is, the probability of defection when the opponent's action is unknown should equal a weighted average of the probabilities for the two cases when the opponent's action is known. Contrary to this classical law, the probability that the player defected in the unknown condition fell far outside the bounds of any weighted average. These findings have been interpreted within the heuristic framework as a failure of consequential reasoning [60]. Quantum theory, however, provides a simple, axiomatic explanation based on the same principles of incompatibility as used in the previous examples [45].

Violations of the law of total probability occur in many other applications, such as in experiments investigating the interference of categorization on subsequent decision-making [49]. In the experiments, participants were shown faces, and on some trials they were asked to categorize them as “good” guys or “bad” guys, and then decide to “attack” or “withdraw” (the categorization-decision condition), and on other trials they only had to decide to “attack” or not without explicitly categorizing a face (the decision-alone condition). They were usually (e.g., 70% of the trials for both conditions) rewarded for attacking bad guys and withdrawing from good guys. As expected, if the person categorized the face as a good guy, then the probability of attacking was low; if the face was categorized as a bad guy, the probability of attacking was much higher. Surprisingly, however, the probability of attacking without categorization (in the decision-alone condition) exceeded the probability to attack when the face had been categorized as a bad guy (in the categorization-decision condition). Once again, this is a violation of the law of total probability, because according to this law,

$$p(Attack) = p(Good\ Guy) \cdot p(Attack|Good\ Guy) + (Bad\ Guy) \cdot p(Attack|Bad\ Guy).$$

Contrary to this law, the experiments found that the probability of attacking for the decision-alone condition falls outside the bounds of any weighted average of the probabilities conditioned on the categories.

Quantum theory once again uses the principle of incompatibility to provide a straightforward account of violations of the law of total probability [45, 49]. As illustrated by Panel c in Figure 3 and using the categorization-decision example, the probability of attacking in the decision-alone condition is determined by the projection from the state vector *S* down to the axes for the attack decision, which produces a probability of .875. If the person first categorizes the face as a bad guy, then the state collapses onto this category to form a new state, which is now aligned with the red “bad guy” axis. The probability of attacking, given the face is categorized as a bad guy, is then determined by projecting from the state vector aligned with the red “bad guy” axis down to the horizontal blue axis representing the attack action, which produces a probability of .603. Thus, the probability to attack (.875) in the decision-along condition is higher than the higher bound of the weighted average (.603).

Critics might argue that quantum models’ ability to account for all these findings relies on being more complex than their traditional counterparts. This criticism has been directly addressed by performing a rigorous quantitative comparison of quantum vs. previously successful traditional, classical decision models using a Bayesian model comparison method [61]. Note that model complexity depends not only on the number of model parameters but also on its functional flexibility [62]. The Bayesian model comparison method evaluates models with respect to their accuracy, parsimony, and robustness. In the model comparison [61], the Bayes factor strongly favored the quantum model over the traditional decision models, suggesting the quantum model is more accurate and robust without being more complex than models based on classical probability theory.

Summary

By means of three simple, intuitive examples, this review illustrates how some seemingly diverse and puzzling cognitive phenomena can be explained by a common set of underlying principles drawn from quantum theory. Of course, it may be possible to construct a traditional model for each of these findings alone—for example, support theory [63] accounts for the conjunction fallacy but not the other two findings. However, the real challenge is to find a single coherent explanation for them all. Figure 3 demonstrated how quantum theory accomplishes this challenge—using simple 2-dimensional "toy" models for ease of exposition. The quantum models that were actually used to account for all the quantitative details related to these empirical findings require more than just two dimensions. However, the principles in the more complicated models are identical to those illustrated in the "toy" examples. Box 5 provides a general summary comparison of Bayesian versus quantum models of cognition.

Concluding remarks

This review began by asserting that there are at least two viable ways to formulate probabilistic models of cognition—one using classical probability theory and another using quantum probability theory. We do not wish to leave the impression that these two theories are in strict conflict. In fact, although their formulations are emphatically different, they do agree in many cases. In particular, there is perfect agreement when all the events under consideration are compatible. The need for the quantum approach only arises when incompatible events are involved, which necessarily imposes a sequential evaluation of the events (see Box 2). This incompatibility produces superposition states of uncertainty that result in violations of some of the important laws of classical probability theory.

As a growing new field, quantum cognition also raises many new questions for cognitive scientists and psychologists to address (see Box 6). One is about rationality. Classical probability theory may provide an upper bound that achieves optimal performance irrespective of computational costs and resource limitations. Quantum models may provide a more realistic bound that performs close to optimal but with fewer computational demands. Consider, for example, a strategic game involving yourself and $n - 1$ other players, and each player can choose one of K actions. If human cognition directly implements classical probability theory (i.e., it treats all events as compatible), then K^n joint probabilities are required to represent your beliefs regarding the actions that could be taken by yourself and the other $n - 1$ players, producing an exponential growth in dimensionality. If instead, human cognition applies a new incompatible perspective to each player, then all of the required probabilities can be assigned by using a single state vector that is evaluated with respect to different bases within a fixed K-dimensional space. Using limited cognitive resources [75], incompatibility provides humans the means for answering an unlimited number of questions, thus promoting parsimony and cognitive economy [76]. However, the use of incompatibility comes at the cost of introducing non-commutativity and sequential effects. Our view is that incompatibility of events provides an effective solution to bounded resources, which is the reason for bounded rationality, and for this reason, we argue that incompatibility is ubiquitous in psychology [77].

Acknowledgement and author notes:

The work was supported by U.S. Air Force Office of Scientific Research (FA 9550-12-1-0397) to Z.W. and J.R.B., and the Australian Research Council (DP0773341, DP1094974) to P.B. The authors contributed equally to this article.

References

[1] Simon, H. A. (1957) *Models of Man: Social and Rational*, Wiley

[2] Gigerenzer, G. and Todd, P. M. (1999) *Simple Heuristics That Make Us Smart*, Oxford University Press

[3] Payne, J. W. et al. (1993) *The Adaptive Decision Maker*, Cambridge University Press

[4] Kahneman, D. et al. (1982) *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press

[5] Savage, L.J. (1954) *The Foundations of Statistics*, John Wiley & Sons

[6] Chater, N. et al. (2006). Probabilistic models of cognition: conceptual foundations. *Trends Cogn. Sci.* 10, 287-291

[7] Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn. Sci.*, 10, 309-318.

[8] Busemeyer, J. R. and Bruza, P. D. (2012) *Quantum Models of Cognition and Decision,* Cambridge University Press

[9] Pothos, E. M. and Busemeyer, J. R. (2013) Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain. Sci.* 36, 255-274

[10] Wang, Z. et al. (2013) The potential of using quantum theory to build models of cognition. *Top. Cogn. Sci.* 5, 672-688

[11] Narens, L. (2015) *Probabilistic Lattices with Applications to Psychology*, World Scientific

[12] Bordley, R. F. (1998) Quantum mechanical and human violations of compound probability

principles: toward a generalized Heisenberg uncertainty principle. *Oper. Res.* 46, 923-926

[13] Aerts, D. and Aerts, S. (1994) Applications of quantum statistics in psychological studies of decision processes. *Found. Sci.* 1, 85-97

[14] Khrennikov, A. Y. (1999) Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social and anomalous phenomena. *Found. Physics* 29, 1065-1098

[15] Atmanspacher, H. et al. (2002) Weak quantum theory: complementarity and entanglement in physics and beyond. *Found. Physics* 32, 379–406

[16] von Neumann, J. (1932/1955) *Mathematical Foundations of Quantum Theory,* Princeton University Press

[17] Rdei, M. and Stoeltzner, M. eds (2001) *John von Neumann and the Foundations of Quantum Physics,* Kluwer Academic

[18] Tourangeau, R. et al. (2000) *The Psychology of Survey Response*, Cambridge University Press

[19] Schuman, H. and Presser, S. (1981) *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*, Academic Press

[20] Bohr, N. (1958) *Atomic Physics and Human Knowledge*, Wiley

[21] Heisenberg, W. (1958) *Physics and Philosophy*, Harper and Row

[22] Griffiths, R. B. (2003) *Consistent Quantum Theory*, Cambridge University Press

[23] Kolmogorov, A. N. (1933/1950). *Foundations of the Theory of Probability*, Chelsea Publishing

[24] Suppes, P. (1966) The probabilistic argument for a nonclassical logic in quantum mechanics. *Philosophy of Sci.* 33, 14–21

[25] Rumelhart, D. E. et al. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, MIT Press

[26] Stewart, T. C. and Eliasmith, C. (2013) Realistic neurons can compute the operations needed by quantum probability theory and other vector symbolic structures. *Behav. Brain Sci.* 36, 307-308

[27] Birkhoff, G. and von Neumann, J. (1936) The logic of quantum mechanics. *Ann. Math.* 37, 823-843

[28] Atmanspacher, H. et al. (2004) Quantum zeno features of bistable perception. *Bio. Cyber.* 90, 33-40

[29] Conte, E. et al. (2009) Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Sys. Info. Dyn.* 16, 1-17

[30] Bruza, P. et al. (2009) Is there something quantum-like about the human mental lexicon? *J. Math. Psy.* 53, 362-377

[31] Brainerd, C. J. et al. (2013) Superposition of episodic memories: overdistribution and quantum models. *Top. Cogn. Sci.* 5, 773-799

[32] Nelson, D. L. et al. (2013) How activation, entanglement, and searching a semantic network contribute to event memory. *Mem. Cogn.* 41, 797-819

[33] Aerts, D. et al. (2013) Concepts and their dynamics: A quantum theoretic modeling of human thought. *Top. Cogn. Sci. 5*, 737-772

[34] Aerts, D. (2009) Quantum structure in cognition. *J. of Math. Psychol.* 53, 314-348

[35] Aerts, D. and Gabora, L. (2005) A theory of concepts and their combinations II: a Hilbert space representation. *Kybernetes* 34, 192-221

[36] Gabora, L. and Aerts, D. (2002) Contextualizing concepts using a mathematical

generalization of the quantum formalism. *J. Exp. Theo. Art. Intel.* 14, 327-358

[37] Wang, Z. et al. (2014) Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl. Acad. Sci. U. S. A.* 111, 9431-9436

[38] White, L. C. et al. (2014) Sometimes it does hurt to ask: the constructive role of articulating impressions. *Cogn.* 133, 48–64

[39] Busemeyer, J. R. et al. (2011). A quantum theoretical explanation for probability judgment errors. *Psychol. Rev.* 118, 193-218

[40] Trueblood, J. S. and Busemeyer, J. R. (2011) A quantum probability explanation for order effects on inference. *Cogn. Sci.* 35, 1518-1552

[41] Busemeyer, J. R. et al. (2015) The conjunction fallacy, confirmation, and quantum theory: Comment on Tentori, Crupi, and Russo (2013). *J. Exp. Psychol. Gen.* 144, 236-243

[42] Asano, M. et al. (2012) Quantum-like generalization of the Bayesian updating scheme for objective and subjective mental uncertainties. *J. Math. Psychol.* 56, 166-175

[43] Trueblood, J. S. and Busemeyer, J. R. (2012) A quantum probability model of causal reasoning. *Front. Psychol.* 3, Article 138

[44] Lambert-Mogiliansky, A. et al. (2009) Type indeterminacy: a model of the KT (Kahneman-Tversky) man. *J. Math. Psychol.* 53, 349-361

[45] Pothos, E. M. and Busemeyer, J. R. (2009) A quantum probability explanation for violations of 'rational' decision theory. *Proc. R. Soc. B* 276, 2171-2178

[46] Yukalov, V. I. and Sornette, D. (2011) Decision theory with prospect interference and entanglement. *Theory and Dec.* 70, 283-328

[47] La Mura, P. (2009) Projective expected utility. *J. Math. Psychol.* 53, 408-414

[48] Busemeyer, J. R. et al. (2006) Quantum dynamics of human decision making. *J. Math.*

Psychol. 50, 220-241

[49] Busemeyer, J. R. et al. (2009) Empirical comparison of Markov and quantum models of decision making. *J. Math. Psychol.* 53, 423-433

[50] Danilov, V. I. and Lambert-Mogiliansky, A. (2010) Expected utility theory under non-classical uncertainty. *Theory Dec.* 68, 25-47

[51] Kvam, P. D. et al. (2014) An empirical test of type-indeterminacy in the Prisoner's Dilemma. In *Quantum Interaction* (Atmanspacher, H. et al., eds) pp. 213-224, Springer

[52] Lambert-Mogiliansky, A. and Busemeyer, J.R. (2012) Quantum type indeterminacy in dynamic decision making: Self-control through identity management. *Games* 3, 97-118

[53] Tversky, A. and Kahneman, D. (1983) Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychol. Rev.* 90, 293-315

[54] Carlson, B.W. and Yates, J.F. (1989) Disjunction errors in qualitative likelihood judgment. *Org. Behav. Human Dec. Proc.* 44, 368-379

[55] Stolarz-Fantino, S. et al. (2003) The conjunction effect: new evidence for robustness. *American J. Psychol.* 116, 15-34

[56] Moore, D.W. (2002) Measuring new types of question-order effects. *Public Opinion Quart.* 66, 80-91

[57] Wang, Z. and Busemeyer, J.R. (2013) A quantum question order model supported by empirical tests of an *a priori* and precise prediction. *Top. Cogn. Sci.* 5, 689-710

[58] Hogarth, R.M. and Einhorn, H.J. (1992) Order effects in belief updating: the belief-adjustment model. *Cogn. Psychol.* 24, 1-55

[59] Hastie, R. and Park, B. (1986) The relationship between memory and judgment depends on whether the judgment task is memory-based or online. *Psychol. Rev.* 93, 258-268

[60] Shafir, E. and Tversky, A. (1992) Thinking through uncertainty: nonconsequential reasoning and choice. *Cogn. Psychol.* 24, 449-474

[61] Busemeyer, J. R. et al. (2015). Bayesian model comparison favors quantum over standard decision theory account of dynamic inconsistency. *Decision* 2, 1-12

[62] Myung, I.J. and Pitt, M.A. (1997) Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Rev.* 4, 79-95

[63] Tversky, A. and Koehler, D. (1994) Support theory: a nonextensional representation of subjective probability. *Psychol. Rev.* 101, 547-567

[64] Dzhafarov, E.N. and Kujala, J.V. (2012) Selectivity in probabilistic causality: Where psychology runs into quantum physics. *J. Math. Psychol.* 56, 54-63

[65] Bruza, P. et al. (2010). The role of non-factorizability in determining “pseudo-classical" non-separability. In *Quantum Informatics for Cognitive, Social, and Semantic Processes* (Bruza, P. et al., eds), pp.26-31. AAAI Press

[66] Dzhafarov, E.N. and Kujala, J.V. (2013) All-possible-couplings approach to measuring probabilistic context. *Plos One* 8, article e61712

[67] Hameroff, S.R. (1998) Quantum computation in brain microtubules? The Penrose-Hameroff ″Orch OR″ model of consciousness. *Philos. Trans. R. Soc. Lond. A* 356, 1869-1896

[68] Schwartz, J.M. et al. (2005) Quantum physics in neuroscience and psychology: a neurophysical model of mind-brain interaction. *Philos. Trans. R. Soc. Lond. B* 360, 1309-1327

[69] Freeman, W.J. and Vitiello, G. (2006) Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. *Phys. Life Rev.* 3, 93-118

[70] Hameroff, S. R. (2013) Quantum mechanical cognition requires quantum brain biology. *Behav. Brain Sci.* 36, 287–288

[71] de Barros, J.A. and Suppes, P. (2009) Quantum mechanics, interference, and the brain. *J. Math. Psychol.* 53, 306-313

[72] beim Graben, P. and Atmanspacher, H. (2006) Complementarity in classical dynamical systems. *Found. Phys.* 36, 291-306

[73] beim Graben, P. et al. (2013) Epistemic entanglement due to non-generating partitions of classical dynamical systems. *Int. J. Theor. Phys.* 52, 723-734

[74] Pothos, E.M. and Busemeyer, J.R. (2014) In search for a standard of rationality. *Front. Psychol*. 5, article 49

[75] Cowan, N. (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87-114

[76] Atmanspacher, H. and Römer, H. (2012) Order effects in sequential measurements of non-commuting psychological observables. *J. Math. Psychol.* 56, 274-280

[77] Khrennikov, A. (2010) *Ubiquitous Quantum Structure: From Psychology to Finance,* Springer

Figure Legends

Figure 1. Cognitive science interpretation of a projector *P*.

The state labeled *S* is a distributed representation across a set of input nodes. The projector *P* is the matrix of weights connecting input nodes to output nodes. The inputs are mapped by the connections to the output nodes to produce a distributed representation of the projection labeled *T*. The probability of an event equals the squared length of the projection, $||T||^2$. This mapping is analogous to a connectionist network, except that the values assigned to the input and output nodes are not restricted to be positive real numbers as required by standard connectionist networks.

Figure 2. A hypothetical example to illustrate the quantum principles of complementarity, uncertainty, and superposition.

Hypothetical example using a three-dimensional vector space to represent three possible answers to the trip enjoyment question. Panel *a* shows the representation of the three answers to the enjoyment question from the self-perspective (in black). Panel *b* shows the representation of these answers from the other-perspective (in black) that is rotated from the self-perspective (in light grey). Event *A* represents the event "you will enjoy the trip," *B* represents "you are neutral about the trip," and event *C* represents "you will not enjoy the trip." The event *A* is represented by the subspace defined as the ray or axis *A* in the left panel; likewise, for events *B* and *C*. Axes *U, V,* and *W* represent these same three events but based on taking your friend's perspective. Within each panel, the blue vector labeled *S* represents the cognitive state of the person who makes the judgments. The event that *A* or *B* (you will enjoy or be neutral about the trip) occurs is represented by the flat plane formed by the *A*, *B* axes. In panel *a*, the vector labeled *T* in the *A, B* plane represents the projection on the subspace for the event *A* or *B* from the self-perspective, and the squared length of the projection *T* is the probability of event *A* or *B*. Similarly, in panel *b*, the vector labeled *R* represents the projection on the subspace *V* from the other-perspective, and the squared length of the projection *R* is the probability of event V (your friend will be neutral about the trip).

Figure 3. Three example applications of a cognitive model based on quantum probability theory.

Using a small set of coherent quantum probability rules, the model explains various empirical findings that are puzzling to classical probability theory. Panel *a* illustrates the application to the conjunction fallacy paradigm [39,41]. The feminist perspective is represented as a two-dimensional vector space where the blue axis *F* corresponds to the decision “Linda is a feminist" and $\neg F$ corresponds to “Linda is not a feminist." The same two-dimensional vector space, but with rotated axes (the red axes), corresponds to the perspective of Linda being a bank teller (*B*), or not ($\neg B$). Initially, the cognitive state of the decision maker is represented by the magenta colored vector *S*, which is suspended between both sets of axes.

Panel *b* illustrates the application to question order effects [37, 57]. The horizontal blue axis in the figure represents the event “yes” to the Clinton question, and the vertical blue axis represents the "no" answer to the question. The two orthogonal red axes are used to represent the answers to the Gore question, which are rotated 45 degrees (in this simple illustration) with respect to the blue axes for the Clinton question. The initial attitude state of the respondent before answering the two questions is represented by the magenta colored vector labeled *S*.

Panel *c* illustrates an application to the categorization-decision paradigm [49]. The horizontal blue axis is used to represent the "attack" action, and the vertical blue axis is used to represent the "withdraw" action. The categorization responses of "good guy" or "bad guy" are represented by the red axes, which are rotated with respect to the blue axes. The magenta colored vector *S* represents the state of the decision maker before making any categorization, which is superposed with respect to categories.

Figure 4. An empirical test of the QQ equality.

An empirical test of the QQ equality using 70 national field experiments in the U.S. (in addition to two lab experiments) [36]. Each point in this figure represents an order effect from one side of the QQ equality plotted as a function of the order effect shown on the other side. The QQ equality predicts that these points should all be on the straight line that has a slope equal to negative one and intercept equal to zero. The observed points fall closely along this line with a correlation equal to $r = -.82$.

Additional Materials (Glossary, Text Boxes, and Outstanding Questions)

Glossary

Bayesian models of cognition. Bayes rule is a simple theorem that follows from the classical probability definition of conditional probability. Suppose $\{H_1, \ldots, H_N\}$ is a set of hypotheses that you wish to evaluate, and D is some data that provide evidence for or against each hypothesis. Then according to the definition of conditional probability (see Box 1), $p(H_i|D) = p(H_i \cap D)/p(D)$. Bayes rule uses the classical definition of joint probability to rewrite the numerator on the right hand of the equation: $p(H_i \cap D) = p(H_i) \cdot p(D|H_i)$; and Bayes rule uses the law of total probability to rewrite the denominator: $\sum_j p(H_j) \cdot p(D|H_j)$. Bayesian models of cognition use these rules to construct models that predict how people make complex inferences from a set of observations [6,7].

Compatibility. Two questions are compatible if they can be answered simultaneously, or even if they are answered sequentially, the order does not matter; two questions are incompatible if they have to be asked sequentially and the order does matter. The principle of complementarity posits that some questions are incompatible, and these incompatible questions provide different perspectives for understanding the world, and these different perspectives are needed for a complete understanding of the world. Classical probability models usually assume *unicity*, which means all events can be described within a single compatible collection of events. In comparison, incompatible events are unique to quantum theory, which does not impose the principle of unicity.

Conjunction fallacy. Classical probability theory usually assumes that events are as subsets of a single sample space. This implies that the probability of an event A can never be less than the probability of the conjunction of A with another event B ("A *and* B"): $p(A) \geq p(A \cap B)$. However, violations of this law of classical probability, called the "conjunction fallacy", have been found in empirical studies. The most well known empirical study is the Linda problem described in the review [e.g., 53], where human subjects rated the conjunction to be more probable. A quantum model has been proposed which explains the conjunction fallacy, along with other paradoxical findings [39, 41].

Contexuality. Constructing a classical probabilistic model involves defining relevant variables, which in turn form the basis of a joint probability distribution over the variables. However, research on entangled quantum systems has taught us that we cannot always assume the existence of joint distributions, and this approach to constructing probabilistic models can fail when applied to the observed data. This failure has come to be known as "contextuality." It refers to the inability to construct the joint distribution over the variables.

Disjunction fallacy. The classical probability theory also implies that the probability of the disjunction of an event A with another event B ("A *or* B") can never be less than the probability of the event A: $p(A) \leq p(A \cup B)$. Violations of this classical probability rule, called "disjunction fallacy," have been found in empirical studies [e.g., 54]. The same quantum cognition model used to explain conjunction fallacy also explains the disjunction fallacy [39, 41].

Dutch book argument (DBA). Decision scientists and probability theorists use the Dutch book argument to show that classical probability theory is a rational theory. The idea originated with Bruno de Finetti, who proposed a game between a bookmaker and a better, where the bookmaker provided stakes for bets that reflected his probability of winning. The better could make a Dutch book against the bookmaker if the bookmaker's stakes for individual bets were chosen in a way that the sum across bets guaranteed that the better would win money and the bookmaker would lose money in every state of the world. If the bookmaker chooses stakes that satisfy an additive measure (items 3,4 in Box 1), then no Dutch book can be made against the bookmaker.

Hilbert space. A Hilbert space is an abstract and complete vector space defined on the complex field and possessing the operation of an inner product (or dot product). It is named after the famous mathematician David Hilbert. It extends vector algebra and calculus from the two-dimensional Euclidean plane and three-dimensional space to spaces with an arbitrary number of dimensions, including spaces of infinite dimensions. A finite Hilbert space is an N-dimensional vector space defined on a field of complex numbers and endowed with an inner product.

Superposition. Superposition is a basic principle of quantum probability theory. Classical probability theory assumes that at any moment, a system is in a definite state with respect to possible states. This definite state can change stochastically across time, but at each moment, the state is still definite, and the system produces a definite sample path. In contrast, quantum probability theory assumes that at any moment, a system is in an indefinite (technically dispersed) superposition state until a measurement is performed on the system. To be in a superposed state means that all possible definite states have the *potential* for being actualized, but only one of them will *become* actual upon measurement. The concept of superposition resonates with the fuzzy, ambiguous, uncertain feelings in many psychological phenomena.

The law of total probability. In classical probability theory, the law of total probability is a fundamental rule relating marginal probabilities to conditional probabilities. It is derived from the distributive axiom of Boolean logic: If $\{A, B, C\}$ are events, then, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. Define $p(A)$, $p(B)$, and $p(\sim B)$ as the marginal probability of event A, B, and ~B, respectively; and define $p(A|B)$ and $p(A|\sim B)$ as the conditional probability of event A conditioned on knowing event B and ~B, respectively. Then the law of total probability is, $p(A) = p(B)p(A|B) + p(\sim B)p(A|\sim B)$. This law provides the foundation for inferences with Bayes networks. In some experiments, $p(A)$ is estimated from one condition, and $p(B)p(A|B) + p(C)p(A|C)$ is estimated from another condition, and violations of this classical law have been found [e.g., 49, 60].

The "sure thing" principle. Savage [5] introduced the "sure thing" principle as a normative principle governing rational decision making. According to this principle, if under the state of the world X, a person prefers action A over B, and if under the complementary state of the world *not*

X, the person also prefers action A over B, then the person should prefer action A over B even when she/he does not know the state of the world. Violations of the sure thing principle have been empirically found [e.g., 60]: When A is preferred over B for each known state of the world, the opposite preference occurs when the state of the world is unknown.

Box 1. Some key axioms, definitions, and theorems of classical probability theory.

1. Events are subsets of a universal set U. Events, such as A and B, are subsets of U.
2. The state of the cognitive system is represented by a function p defined on the subsets in U and the probability of an event A equals $p(A)$.
3. $p(A) \geq 0$, and $p(U) = 1$.
4. If $A \cap B = \emptyset$, then $p(A \cup B) = p(A) + p(B)$.
5. The probability of event B given A equals $\frac{p(A \cap B)}{p(\boldsymbol{A})}$.
6. Law of total probability: $p(B) = p(A \cap B) + p(\sim A \cap B)$.

Box 2. Some key axioms, definitions, and theorems of quantum probability theory.

1. Events are subspaces of a Hilbert space $\mathcal{H}$. Events, such as A and B, correspond to subspaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively of $\mathcal{H}$. Associated with these subspaces are projectors P_A and P_B.
2. If their projectors are commutative, that is, $P_A P_B = P_B P_A$, then the events A and B are compatible. Otherwise, they are incompatible.
3. The state of the cognitive system is represented by a unit length vector S in the vector space, and the probability of event A equals $\|P_A \cdot S\|^2$.
4. $\|P_A \cdot S\|^2 \geq 0$ and $\|P_{\mathcal{H}} \cdot S\|^2 = 1$.
5. If $P_A P_B = 0$, then $\| (P_A + P_B) \cdot S\|^2 = \|P_A \cdot S\|^2 + \|P_B \cdot S\|^2$.
6. Probability of event B given A equals $\frac{\|P_B \cdot P_A \cdot S\|^2}{\|P_A \cdot S\|^2}$.
7. Violation of the law of total probability: $\|P_B \cdot S\|^2 \neq \|P_B \cdot P_A \cdot S\|^2 + \|P_B \cdot P_{\sim A} \cdot S\|^2$.

Box 3. The Linda problem.

The Linda problem is a typical example showing conjunction fallacy [53]. In the experiments, participants are given the following scenario:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Which of the following is more probable?
> (a) Linda is a bank teller.
> (b) Linda is a bank teller and is active in the feminist movement.

According to classical probability theory, the probability of the conjunction of events (b) can never exceed the probability of one of its constituent events (a). However, many studies have robustly shown that human participants consistently rate option (b) as more probable than (a).

Box 4. Payoff matrix for the prisoner's dilemma game.

In a typical prisoner's dilemma game, the payoff matrix is set up in the way that no matter what move your opponent makes, you are better off to defect; the same is true for your opponent. The following table shows an example.

	the opponent defects	*the opponent cooperates*
you defect	You get 10, and the opponent gets 10;	You get 25, and the opponent gets 5;
you cooperate	You get 5, and the opponent gets 25;	You get 10, and the opponent gets 10.

Box 5. Comparison of Bayesian and quantum models of cognition

Bayesian models of cognition [6,7] have been established for a longer period of time.	Quantum models of cognition [8-10] are recently developed.
Bayesian models are based on a probability theory that is more conventional and more familiar to most researchers.	Quantum models are based on a probability theory that is unconventional and unfamiliar to many researchers, especially in the fields of social and behavioral sciences.
Bayesian models are derived from a set of axioms that may be too restrictive for describing all of human judgments and decision making.	Quantum theory is based on a more relaxed set of axioms that may better reflect human reasoning and decision making under uncertainty.
Bayesian models have many successful applications in areas such as concept learning and causal reasoning.	Although quantum models have been applied to many different cognitive phenomena, they have mainly focused on problems found in judgment and decision research that are difficult for Bayesian models to explain.
Stronger than usual evidence is required to encourage researchers to accept the newer ideas of quantum theory over the more established Bayesian models.	Tests of *a priori* predictions and quantitative model comparisons are beginning to provide strong evidence needed to justify quantum models of cognition.
It may be possible to develop a classical model to account for all of the puzzling findings that are used to support quantum cognition.	At this point in time, no such classical model has been proposed. If such a classical model is proposed, then rigorous model comparisons are required to compare the classical and quantum models.
Bayesian models may be considered the normative standard without regard to computational costs and resource limitations.	Quantum models may be considered the prescriptive standard that works effectively, while at the same time obeying the constraints of bounded rationality.

Box 6. Outstanding questions.

- *What about dynamics of cognition and decision? In the review, we have only presented the structural component of quantum theory. Can the Schrödinger wave equation be used to describe the dynamic processes that lead to decisions? If so, how do quantum dynamics differ from classical dynamics based on Markov processes [48]?*

- *Will contextuality create a major shift in how probabilistic models are viewed within cognitive science? Constructing classical probabilistic models involves defining relevant variables, which in turn form the basis of a joint probability distribution over the variables. Does this joint distribution always exist? Pioneering researchers investigating contextuality are beginning to question this assumption [32, 64-66].*

- *If the quantum cognition research program does not rely on the assumption that the brain is some type of quantum computer, then how does the brain accomplish quantum-like behavior? Scientists are still far from understanding the way mental states emerge from the neural substrates. Is quantum physics related here in a significant way [67-70]? Or, can quantum-like effects emerge from a classical neural network [71-73]?*

- *How to define rationality? Are quantum probability models rational or irrational? Quantum probability, like classical probability, is an axiomatic system for assigning probabilities to events, but the two systems rely on different axioms. Can either system of axioms legitimately lay claim to be the norm for prescribing rational behavior in reasoning, inference, and decision-making? Do quantum probabilities satisfy the Dutch book argument [74]?*

- *If the world that we as humans experience is the macro-level world in which objects behave to a large extent according to the rules of classical physics, then why would evolution result in a reasoning system based on quantum rather than classical probability? Does everything in the macro-level world that we experience behave according to classical probability theory? What about human beings as well as animals with which we interact in this macro-world? Do conscious mental events obey rules of classical physics?*