Quantum cognition: a new theoretical approach to psychology

Peter D. Bruza1, Zheng Wang2, and Jerome R. Busemeyer3

1Information Systems School, GPO Box 2434, Queensland University of Technology, Brisbane 4001, Australia
2School of Communication, 3145 Derby Hall, 154 North Oval Mall, The Ohio State University, Columbus, OH 43210, USA
3Department of Psychological and Brain Sciences, 1101 East 10th Street, Indiana University, Bloomington, IN 47805, USA

What type 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 have 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 versus 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 crystallize 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, ‘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 do 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 because, 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 crucial 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 some 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 crucial 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 physics to the field of human cognition [12–15]. They did not proceed from the assumption that there is something quantum-like taking place 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 1930s. 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, and expressing the combined event ‘A and B’ is therefore the same as stating ‘B and A’. In other words, the order of the events does not matter. By 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 ‘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 because 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 made. In principle, these decisions can be evaluated simultaneously, and evaluating one does not interfere with what we think about the other, and the same outcomes are therefore 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 types of events, compatible and incompatible, 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 must be considered 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 best-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, together with the uncertainty and superposition principles, leads to key differences between classical and quantum theories of probability [22]. These principles, we will argue, are crucial for understanding the limitations operating 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 consider 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, A ∩ 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, A ∪ B, which in our example represents the disjunction that you think you or 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 A ∩ B is an event, and so is A ∪ B. Recall that this logic of events is commutative, that is, A ∩ B = B ∩ A. Another important property of this logic is the distributive axiom:

\[
A = A \cap (B \cup \sim B) = (A \cap B) \cup (A \cap \sim B)
\]

where  \(\sim B\) 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

---

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

(i) Events are subsets of a universal set \(U\). Events, such as \(A\) and \(B\), are subsets of \(U\).

(ii) 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 \(p(A)\).

(iii) \(p(A) \geq 0\), and \(p(U) = 1\).

(iv) If \(A \cap B = \emptyset\), then \(p(A \cup B) = p(A) + p(B)\).

(v) The probability of event \(B\) given \(A\) equals \(p(B|A) = \frac{p(A \cap B)}{p(A)}\).

(vi) 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

(i) 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 $P_B$.

(ii) 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.

(iii) The state of the cognitive system is represented by a unit length vector $\mathbf{S}$ in the vector space, and the probability of event $A$ equals $||P_A \cdot \mathbf{S}||^2$.

(iv) $||P_A \cdot \mathbf{S}||^2 = 0$ and $||P_B \cdot \mathbf{S}||^2 = 0$.

(v) If $P_A P_B = 0$, then $||P_A + P_B \cdot \mathbf{S}||^2 = ||P_A \cdot \mathbf{S}||^2 + ||P_B \cdot \mathbf{S}||^2$.

(vi) The probability of event $B$ given $A$ equals $\frac{||P_B P_A \cdot \mathbf{S}||^2}{||P_A \cdot \mathbf{S}||^2}$.

(vii) Violation of the law of total probability: $||P_B \cdot \mathbf{S}||^2 \neq ||P_B \cdot P_A \cdot \mathbf{S}||^2 + ||P_B \cdot P_A \cdot \mathbf{S}||^2$.

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, $\mathbf{S}$, to assign these probabilities as described in Box 2.

Mathematically, a vector is an arrow lying within a vector space. The state vector, $\mathbf{S}$, 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 $\mathbf{S}$ 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] (Figure 1). This mapping process is termed 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 $\mathbf{S}$ 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 $\mathbf{S}$ onto the plane spanned by the $A$, $B$ axes, represented by the blue vector $T$. Projection is akin to 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 2A) – 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 0.833 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 $\mathbf{S}$ onto the axis $C$ (which equals 0.167 in the example).

Now consider the enjoyment of the trip from the perspective of your friend by referring to panel B in 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 (Figure 2B). 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 $\mathbf{S}$ 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 $\mathbf{S}$ onto ray $V$ (which equals 0.812 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 because 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 $\mathbf{S}$ 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 0.375. If instead we first project the state $\mathbf{S}$ 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 0.406. In other words, reversing the order of projections alters the final probability of two
decisions made 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 type 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 algorithm-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 $A$ and $B$ are incompatible, then the conjunction of events: $A \land B$, cannot be defined because 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 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 $\neg A$ 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 already 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
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?

(i) Linda is a bank teller.
(ii) Linda is a feminist and bank teller.

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

deivate substantially from what would be considered normatively correct according to classical logic and probability theory [39]. A popular example is ‘the Linda problem’. As described 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 Figure 3A, 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 brown B axis (B stands for bank teller). Observe that the projection is short, indicating a low probability (0.024 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 ~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 both 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 0.093 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 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 USA concerning public opinions on the characters of Bill Clinton and Al Gore [56]. Half of the 1002 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 Figure 3B, 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 brown Gore axis, which produces a low probability (0.173 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 brown Gore axis, and then projecting the result down to the blue Clinton axis, which produces a higher probability (0.487 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, known as the QQ (quantum question order) equality [37,57]:

\[ p(C \text{ and then } \sim G) = p(\sim G \text{ and then } C) \equiv \sim p(\sim C \text{ and then } G) = p(G \text{ and then } \sim C) \]

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] (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 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 (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 prisoner’s dilemma (PD) game [60]. In the typical PD game (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

---

**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.

<table>
<thead>
<tr>
<th></th>
<th>The opponent defects</th>
<th>The opponent cooperates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>You defect</strong></td>
<td>You get 10, and the</td>
<td>You get 25, and the</td>
</tr>
<tr>
<td></td>
<td>opponent gets 10</td>
<td>opponent gets 5</td>
</tr>
<tr>
<td><strong>You cooperate</strong></td>
<td>You get 5, and the</td>
<td>You get 20, and the</td>
</tr>
<tr>
<td></td>
<td>opponent gets 25</td>
<td>opponent gets 20</td>
</tr>
</tbody>
</table>
informed a player of the opponent’s move before acting. What happened was 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,

\[
p(\text{Player Defects}) = p(\text{Opp Defects}) \cdot p(\text{Player Defects} | \text{Opp Defects}) + p(\text{Opp Cooperates}) \cdot p(\text{Player Defects} | \text{Opp Cooperates}).
\]

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. 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:

\[
p(\text{Attack}) = p(\text{Good Guy}) \cdot p(\text{Attack} | \text{Good Guy})
+ (\text{Bad Guy}) \cdot p(\text{Attack} | \text{Bad Guy}).
\]

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.
Box 5. Comparison of Bayesian and quantum models of cognition

<table>
<thead>
<tr>
<th>Bayesian models of cognition [6,7] have been established for a longer period of time.</th>
<th>Quantum models of cognition [8–10] are a recent development.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian models are based on a probability theory that is more conventional and more familiar to most researchers.</td>
<td>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.</td>
</tr>
<tr>
<td>Bayesian models are derived from a set of axioms that may be too restrictive for describing all of human judgments and decision making.</td>
<td>Quantum theory is based on a more relaxed set of axioms.</td>
</tr>
<tr>
<td>Bayesian models have many successful applications in areas such as concept learning and causal reasoning.</td>
<td>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.</td>
</tr>
<tr>
<td>It may be possible to develop a conventional model to account for the puzzling findings that are used to support quantum cognition.</td>
<td>At this point in time, no such conventional model has been proposed. If such a conventional model is proposed, then rigorous model comparisons are required to compare the conventional and quantum models.</td>
</tr>
<tr>
<td>Bayesian models may be considered as the normative standard without regard to computational costs and resource limitations.</td>
<td>Quantum models may be considered to be the prescriptive standard that works effectively, while at the same time obeying the constraints of bounded rationality.</td>
</tr>
</tbody>
</table>

of the law of total probability [45,49]. As illustrated in Figure 3C, 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 onto the axes for the attack decision, which produces a probability of 0.875. 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 brown ‘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 brown ‘bad guy’ axis down to the horizontal blue axis representing the attack action, which produces a probability of 0.603. Thus, the probability to attack (0.875) in the decision-alone condition is higher than the higher bound of the weighted average (0.603).

Critics might argue that the ability of quantum models 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 versus 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 2D ‘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 only 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 (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 (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
**Box 6. Outstanding questions.**

- What about dynamics of cognition and decision? In this 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 decision? 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 [32,67–69].
- If the quantum cognition research program does not rely on the assumption that the brain is some type of quantum computer, how then 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 [70–73]? Can quantum-like effects emerge from a classical neural network [74–76]?
- 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 [77]?
- 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 the animals with which we interact in this macro-world? Do conscious mental events obey the rules of classical physics?

different bases within a fixed $K$-dimensional space. Using limited cognitive resources [64], incompatibility provides humans the means for answering an unlimited number of questions, thus promoting parsimony and cognitive economy [65]. However, the use of incompatibility comes at the cost of introducing non-commutativity and sequential effects. Our view is that incompatibility comes at the effective solution to bounded resources, which is the reason for bounded rationality [66].

Acknowledgments

The work was supported by US Air Force Office of Scientific Research (FA 9550-12-1-0397) to Z.W. and J.R.B., and the Australian Research Council (DP1073341, DP1094974) to P.B. We thank Marius Usher for comments. The authors contributed equally to this article.

References

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