1
Introduction

Time flies by in a flash and already over 10 years have passed since our first edition. During this time, work in quantum cognition has entered a new stage. We discussed many of the relevant issues and ideas in the first edition of this book – the "first movement" of our "composition." Although we can't keep up with the superluminal pace of quantum entanglement, we can follow the traces left from quantum walking and pluck the hidden strings that play this "second movement" of our "composition."

Quantum mechanics and human behavior are two fields that most people would think are unrelated. For more than two decades, however, scientists have been exploring and clarifying connections between the two fields: Theories in both fields aim to predict how indeterministic systems that are sensitive to measurement will behave in the future. Their difference is that one field aims to understand the nature of the material world through physical processes, while the other aims to understand the nature of our mental world through cognitive processes. Quantum mechanics was originally conceived to explain what seemed to be puzzling behavior at the subatomic level. Likewise, quantum cognition was inspired by the need to account for puzzling behavior at the human level. Classical probability and decision theory are often used to predict how people make inferences and choices from the information they are provided. But there are many manifestations of human behavior that are "contrary to rationality" and so these predictions often fail, sometimes strikingly so. Quantum probability theory turns out to provide robust explanations why these failures occur.

Quantum cognition is a steadily growing new approach to building computational models of cognition and decision based on principles from quantum probability, dynamics, and information processing theory. It is an interdisciplinary field involving researchers from physics, computer science, psychology, social science, and philosophy. Models of quantum cognition need to

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

be distinguished from models based on quantum physics: The former uses only the abstract mathematical principles of the latter, without the physics. For example, a quantum model of cognition might employ a Schrödinger equation that involves a dynamic parameter analogous to a Planck constant, but it certainly won't be the same numerical value as the Planck constant! In other words, quantum cognition is an application of the conceptual framework and formalism of quantum theory to human behavior. It is actually not uncommon for mathematics that was originally developed for application to the physical world to migrate outside of physics. For example, classical diffusion models were originally developed to describe the Brownian motion of molecules in a liquid, but have since been applied outside of physics to finance, disease epidemics, cognitive and neural decision models, and many other fields. The same is happening now with the mathematics from quantum theory: Applications have appeared in psychology (e.g., this book), linguistics [e.g., Heunen et al., 2013], social science [e.g., Bagarello, 2019; Haven and Khrennikov, 2013; Wendt, 2015], finance [e.g., Baaquie, 2004], artificial intelligence [e.g., Wichert, 2014], information retrieval [e.g., Melucci, 2015; Van Rijsbergen, 2004], and engineering [e.g., Dong et al., 2010; Schuld et al., 2014].

1.1 Why Quantum Cognition?

A reader new to this field may wonder: What is quantum about cognition? What makes this a viable approach to understanding human behavior? Quantum physics was developed to describe and predict the behavior of minuscule particles from the subatomic world like photons and electrons. In contrast, humans deal with a macro-level "classical" world, such as, for example, coin flips and billiard balls. Predicting coin flipping behavior only requires classical probability theory, and predicting the motion of colliding billiard balls only requires classical dynamic theory. It is well understood that the behavior of billiard balls can be described by classical physics, but what about the behavior of the billiard players? Classical probability might describe coin flipping, but maybe not a cat dodging an angry dog. Still the fundamental question might persist: What could human behavior possibly have in common with the behavior of a subatomic particle such as an electron? What justifiable reasons are there for considering a quantum approach to cognition? There are a number of answers to these very important questions.

1.1.1 Psychological Reasons

The probabilities generated by a system (e.g., an electron or a person) depend on the state of the system. According to classical theory, it is only the lack of knowledge of the exact state of the system that prevents a deterministic model of behavior. Probabilities arise from this lack of knowledge. This kind of uncertainty is called *epistemic* uncertainty. For example, if we close our eyes and spin a classical spinner, like a roulette wheel, then immediately before we observe it, the spinner is either definitely pointing either more upward or more downward, but not both. Before we look, we can only assign probabilities to each event because of our ignorance of its definite state. However, if the spinner is pointing up just before we observe it, and then we look, we will certainly see what existed (the spinner pointing up) before we looked at it. Similarly, if a juror is following a classical inference process, like Bayes' rule, then at some moment during the trial she has a probability favoring guilty (a probability greater than equally likely) or a probability favoring not guilty (a probability less than equally likely) but not both. Before a judge asks the juror, the prosecutor can only predict the verdict with some probability because he is ignorant of the juror's state. If the juror is favoring guilty just before the judge asks for a verdict, and the judge asks for a verdict, the prosecutor will certainly hear the guilty answer that existed before the judge asked.

According to quantum theory, a system can be in an indefinite state, called a *superposition* state, such that several outcomes have the potential to be realized by a measurement at the same moment. No definite state exists before the measurement. Instead, the measurement creates an observed outcome with some probability, and this probability cannot always be driven to zero by additional knowledge of conditions. This kind of uncertainty is called *ontic* uncertainty. For example, just before we observe it, the spin of an electron can be superposed between spin-up and spin-down directions at that moment: If we observe it at that moment, we could see either up or down, with associated probabilities. These probabilities are not due to our ignorance of a definite spin state, because no definite up or down state existed before we looked. In other words, there is no underlying fact of the matter. Even if we had all information available, we would not be able to determine the definite spin-state of the electron. In quantum physics, the nature of this uncertainty is referred to as *indeterminacy*. The uncertainty is intrinsic to the electron itself, not what we can know about it.

Likewise, if the juror is following a quantum inference process, then before a juror makes up his or her mind, the juror can be superposed between a belief favoring guilty (a belief greater than equally likely) and a belief favoring not guilty (a belief less than equally likely) at the same moment. If the juror is asked for a verdict, the judge could receive a "guilty" verdict or a "not guilty" verdict with associated probabilities. Again these probabilities are not due to ignorance about a definite cognitive state in the juror's mind, because no such definite state existed before the verdict was requested. These probabilities are intrinsic to the juror [see Colyvan, 2004]. As aptly described by Aerts and de Bianchi,

> if the quantum approach to cognition works so well, it is because both the "microscopic layer" of our physical reality, populated by so called quantum "particles," and the "cognitive layer" of our mental reality, populated by conceptual entities, are realms of genuine "potentialities," not of the type of a "lack of knowledge of actualities." [Aerts and de Bianchi, 2015, p. 53]

The formal representation of a superposition state is presented later in Chapter 2.

Sensitivity to measurement is another key property that electrons and humans share. For example, an electron has no definite spin direction before it is measured. Rather, the measurement of the spin *creates* a definite spin direction from the indefinite superposition (see sections 1–5 of Peres [1998] for a discussion of this point). Measuring the spin of an electron in the vertical direction and finding an outcome of spin-up reduces its state from a superposition to a definite state consistent with spin-up. If the state of spin is measured again immediately after, then the outcome is certain to be spin-up. Likewise, the juror is not in a definite cognitive decision state corresponding to the judgment 'guilty' or 'not guilty' before he has made up his mind. Requesting a final verdict *creates* a definite judgment from the underlying indeterminate state. Deciding that a defendant is guilty changes the cognitive state of the juror from an indeterminate state of superposition into a definite state corresponding with a verdict of guilty. If the verdict is requested again immediately after, the outcome is certain to be guilty; however, this effect may not last long because of subsequent dynamic evolution of the cognitive decision state. The creation of a definite state from the indefinite by means of measurement changes the nature of both electrons and humans. This change from indefinite to definite following measurement is called the "collapse" of the superposition state. The word "collapse" is in quotes because the issue about what is collapsing is controversial, which we try to address in Chapter 2.

Some might argue that the reduction in state following a measurement is nothing more than computing a classical conditional probability [Marinoff, 1993]. For example, the probability of rolling a pair of dice and getting a sum greater than 5 is much higher after we observe that the first die turns out to be a 4; the probability that we think a juror will assign a life sentence

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

will be different if we observe that the juror thinks the defendant is guilty. In some ways, this argument is correct, because the "collapse" forms a new updated state conditioned on the observation. However, the reduction that occurs in quantum theory can be more complex than in classical theory because sensitivity to measurement produces what are called *interference* effects [Feynman et al., 1965; Peres, 1998].

In classical theory, the probability of the event "sum of pair is greater than 5" must equal the total probability of "first die is 4 and the sum of pair is greater than 5" or "first die is not 4 and sum of pair is greater than 5." Likewise, if the events "the defendant is guilty" and "the punishment is life imprisonment" are two events in a common classical probability space, then the probability of the event "the punishment is life imprisonment" must equal the total probability of "the defendant is guilty and the punishment is life imprisonment" or "the defendant is not guilty and the punishment is life imprisonment."

In quantum theory, sensitivity to measurements can result in interference effects, which appear to be violations of total probability. For example, consider the effect of measuring spin in the horizontal direction before the vertical direction, as compared to only measuring the vertical direction. The probability that an electron is found to be spin-up when measuring only the vertical direction differs from the total probability that the electron is found to be "spin-left and then spin-up" or "spin-right and then spin-up." Apparently, the event "spin-up" measured alone is not the same as the event "spin-left and then spin-up" or "spin-right and then spin-up," producing an apparent violation of the distributive axiom and hence a violation of the law of total probability. Likewise, the probability of deciding "the punishment is life imprisonment" may differ from the total probability of deciding "the defendant is guilty and then the punishment is life imprisonment" or "the defendant is not guilty and then the punishment is life imprisonment." Once again, the event "the punishment is life imprisonment" is not the same as the event "the defendant is guilty and then the punishment is life imprisonment" or "the defendant is not guilty and then the punishment is life imprisonment." Measurement of a first event changes the nature of a second event as compared to measurement of the latter alone. Interference effects of measurement are very common in human judgments and quantum theory provides a natural way to represent these effects [Khrennikov, 2010]. Interference effects appear in many chapters of this book, especially in Chapters 4 and 5.

A third property that electrons and humans share is called *complementarity*. Actually, Bohr's famous principle of complementarity, which he formulated for physics, may have been originated in a psychological form by William James (James [1890]; see Blutner and beim Graben [2016] for a discussion),

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

and quantum cognition has brought it back to psychology. Bohr's idea was that different measurement conditions are complementary if they are mutually exclusive, but they are all necessary for a comprehensive understanding of nature [Plotnitsky, 2012]. For example, it is not possible to arrange magnets to measure electron spin simultaneously in the up-down vertical direction and in the left-right horizontal direction; instead, they have to be measured sequentially. James [1890] had a different idea that mental thoughts are complementary if they are not simultaneously accessible to the person, but they share knowledge. For example, judgments of guilt and punishment are never made simultaneously, and instead judgments of punishment naturally follow judgments of guilt. Importantly, when events have to be measured sequentially, the sequence can change the results because of sensitivity to measurement. For example, measuring an electron in the up-down vertical direction and then in the left-right horizontal direction produces different results than the opposite order. Similarly, judging guilt before punishment may produce different results than when these judgments are made in the reverse order. In both cases, the first measurement changes the state, which prepares a new context for the second measurement. Of course, some pairs of measurements are sensitive to order and some are not. For example judging something complex and uncertain, such as guilt and punishment of a defendant, may depend on order; but judging other characteristics, such as the gender and height of the defendant, may not.

If a pair of measurements are order dependent, then they are called *incompatible*; if they are not, then they are called *compatible*. If all measurements were compatible, then there would be no difference between quantum and classical probabilities. Question order effects are discussed in more depth in Chapters 4 and 5. Incompatibility and its relationship with indeterminacy is covered in more detail in Chapter 2.

1.1.2 Contextual Reasons

The principle of unicity (see Griffiths, 2003, chapter 27) states that there is a unique exhaustive description which contains all events. In other words, there is a single sample space of points from which all events can be composed. However, the existence of incompatible measurements makes this principle break down so that it is not possible to fit all the events into a single sample space. Essentially, events produced by incompatible measurements require separate sample spaces and separate probability distributions. In this setting, a quantum phenomenon known as *contextuality* can be determined.

Suppose variable A is a yes/no question such as "Do you think the social and economic state of country A is in good shape?" and X is another yes/no

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

question such as "Do you think President X of country A is doing a good job?" Let A denote the measurement of variable A alone, AX denote the measurement of variable A followed by variable X, and XA the opposite order. Then A, AX, and XA form three different measurement contexts. If the variables A, X are compatible, a single two-way joint distribution for AX can account for all of the probabilities of events from all three measurement contexts. However, if A, X are incompatible, so that there are interference effects and order effects, then a single two-way distribution cannot represent the three contexts, and the three distributions must be kept separate.

Different measurement contexts can also be formed by measuring different combinations of variables. For example, suppose we are investigating four binary psychological variables, A, B, X, Y, where A and X are the same as before, B is about country B, and Y is about the president of country B. Then consider four measurement contexts, AX, AY, BX, BY, where for example BX refers to the measurement of variable B and then variable X. We can form a separate 2×2 classical probability distribution for each context to produce a collection of four tables, such as the hypothetical results illustrated in Table 1.1. Even so, we might ask: Is it possible to reconstruct all four separate distributions using a single 2^4-way joint distribution of the four binary variables A, B, X, Y, allowing arbitrary dependencies? It turns out that this may not be possible for several reasons. One reason is that the marginal distributions may be inconsistent. For example, the marginal distribution of variable X in the context AX may be different from the marginal distribution of X in the context BX. Now further suppose that the marginals are all consistent. Can we then reconstruct the four separate distributions, each corresponding to a measurement context, from a single 2^4-way joint distribution (and note that we are allowing any arbitrary dependencies among the four variables)? The answer may still be no, but for a more subtle reason. The four correlations produced by the four distributions could violate a property called the Clauser–Horn–Shimony–Holt (CHSH) inequality, which is required to achieve this reconstruction (discussed later in Chapter 10). The four example distributions shown in Table 1.1 actually violate the CHSH inequality, and so there is no single four-way joint distribution that can reproduce these four tables. Once again, a separate distribution must be used to describe each table.

For larger numbers of measurement contexts, even more constraints must be satisfied, and Dzhafarov and Kujala [2012] identify the general conditions needed to construct a single joint distribution for any collection of measurement contexts. The inability to construct a single joint distribution is seen as the signature of contextuality. Contextuality is a subtle notion that influences

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

Table 1.1 *Numerical example of four two-way tables produced by four contexts**

	$X = y$	$X = n$		$Y = y$	$Y = n$
$A = y$	0.271	0.175	$A = y$	0.115	0.331
$A = n$	0.084	0.469	$A = n$	0.269	0.285
	$X = y$	$X = n$		$Y = y$	$Y = n$
$B = y$	0.335	0.035	$B = y$	0.296	0.073
$B = n$	0.021	0.610	$B = n$	0.088	0.543

*There are some slight rounding errors in the table.

how we must view the properties of the cognitive phenomenon being studied. It is covered in more detail in Chapters 10 and 13.

Quantum probability theory was specifically created to be a contextual theory that can account for the effects of measurement context. A superposition state is used to account for interference effects, such as finding the marginal probability of X in the context AX to be different than the probability of X in the context of XA. Additionally, quantum theory includes the important concept of an *entangled* superposition state to account for deeper contextual effects, such as violations of the CHSH inequality. An entangled superposition state that represents the AX context cannot be decomposed into two separate states, one for variable A and another for X; instead there are interdependencies so that observing the outcome of a measurement of A now changes the probabilities for X. Of course, classical probability theory also allows for dependencies between the variables, but these dependencies must satisfy the CHSH inequality. When entangled states are combined with incompatible measurements (e.g., suppose in our example above, the variables X, Y are incompatible), then quantum theory can provide an elegant account of violations of the CHSH inequality. One then might ask: Is quantum probability empirically testable? In fact, quantum probability must satisfy another inequality, the Tsirelson inequality.

Entangled states are useful for conceptual combinations that are not semantically compositional [Bruza et al., 2015b]. For example, the semantics of the conceptual combination BLACK CAT can be argued as being compositional due to the non-empty intersection of black objects with the set of objects that are cats. In contrast, the intersective semantics of ASTRONAUT PEN are empty, and yet humans can attribute semantics to this combination. How quantum cognition can furnish semantics to language is covered in Chapter 11.

Of course quantum probability is not the only way to account for context effects. However, it provides a general and principled way to do this, rather than relying on ad hoc and specialized assumptions.

1.2 Two Challenges for Quantum Cognition

Often the audience in our talks, or the reviewers of our articles, ask two important questions that challenge the quantum cognition research program. One question asks: Why would biological evolution produce a cognitive system that is quantum-like? The second question concerns the prospects of a neurophysiological basis for quantum-cognitive processing.

1.2.1 Why Would Evolution Pick Quantum Reasoning?

If the *physical world* that we encounter at the macroscopic level is essentially classical, why would evolution generate a cognitive system that uses quantum rather than classical probability? One reason is that perhaps our *mental world* is not adequately described by a classical view. Khrennikov [2007] and Blutner and beim Graben [2016] have both proposed that perhaps the neural system is a classical (deterministic) dynamical system that operates on a high-dimensional continuous state space, called the *micro-state* space. However, the key argument is that a person's *mental experiences* are provided by macroscopic (global brain) measurements of billions of micro-states, which provide a coarse-graining of the micro-states into "macro-states." Information is lost and the state of the microsystem becomes uncertain. Different macroscopic mental measurements can be incompatible, generating different but overlapping Boolean algebras of experienced events [beim Graben and Atmanspacher, 2006]. A collection of different but overlapping Boolean algebras is called a partial Boolean algebra. Assigning probabilities to events that form a partial Boolean algebra is problematic for a single classical probability distribution that relies on unicity. Quantum probability is ideally suited for assigning probabilities to a partial Boolean algebra of events. Thus, the challenge of dealing with a mental world that generates a partial rather than a complete Boolean algebra of experienced events may have prompted the evolution of a quantum probability reasoning system.

Some readers might still not be convinced. Although problematic for a single classical probability distribution, assigning probabilities to a partial Boolean algebra does not necessarily *require* quantum probability, because other generalized probability theories could also apply. So some additional reason is needed to motivate quantum probabilities.

A second reason is based on the idea that quantum probability theory provides more parsimonious (less complex) descriptions than classical joint probability models [Atmanspacher and Römer, 2012]. Many believe that the mind strives to be rational within the limits of its cognitive resources. One popular approach to rational reasoning is Bayesian reasoning, but this approach encounters serious tractability problems. The dimension of a classical joint probability space grows exponentially as the number of variables increases. Consequently, resource limitations of cognition require various simplifications, such as for example, using Bayesian networks that impose strong conditional independence assumptions. This is but one way to be rational within bounded cognitive resources; quantum probability theory provides an alternative to meet the resource constraints for rational reasoning under uncertainty. The dimension of the quantum probability space does not increase exponentially with increasing number of variables. Why is this? As we discuss in the next section, quantum probability defines variables as operators acting on a vector space. (Most computational neural network models actually assume a system operating on a vector space.) The advantage of using a vector space representation is that different variables can be represented by changing the basis (rotating the axes) used to describe them within the *same* vector space. There is an infinite number of ways to select a basis within a fixed, finite-dimensional vector space, which can then provide an infinite number of ways to describe variables within a limited cognitive resource. An example aims to illustrate this important point.

Consider a game with two players, in which each player has three choices of move. When planning a move, each player needs to estimate the probability of the move of the opponent and then consider the probability for his/her own move. According to a Bayesian probability model, this requires forming $3^2 = 9$ joint probabilities that each of two players takes one of three actions. If there are n players, then a Bayesian model requires 3^n joint probabilities, producing an exponential growth in probabilities. In contrast, according to the quantum approach, the state of the three actions by each player could be represented by a vector in a three-dimensional vector space. The probabilities assigned to different players can be obtained by "rotating" the basis used to describe the vector within the same three-dimensional space. In this way, n players are described by n different bases within the same three-dimensional space.

There is, however, a cognitive price to be paid by representing different variables using different bases. Changing the basis used to describe two

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

variables makes the two measurements incompatible. It is not possible to evaluate two incompatible events simultaneously, because changing the basis must be done sequentially by rotating from one basis to another. Furthermore, Heisenberg's famous *uncertainty principle* is a consequence of this incompatibility. For example, in our *n*-person game example, changing the basis used to represent the beliefs about the players implies that it is not possible to be certain about the moves of all players simultaneously. Increasing certainty about the move of one player (rotating to agree with one basis) implies increasing uncertainty (rotating away from another basis) about others. In quantum physics, the uncertainty principle is a consequence of the nature of incompatible observables such as position and momentum. However, for psychology, it may be a consequence of the need to represent an infinite number of variables within a mind constrained by limited cognitive resources [Blutner and beim Graben, 2016]. The uncertainty principle for quantum cognition models is derived in Chapter 2.

1.2.2 Does Quantum Cognition Imply a Quantum Brain?

The quantum brain hypothesis asserts that the brain actually uses quantum physical processes to perform significant cognitive operations [e.g., Hammeroff, 1998, 2007]. We do know that the rods in the human eye can detect a single photon, and therefore some quantum computation is happening at least at very low levels of human information processing [Hecht et al., 1942]. But many have argued strongly against the quantum brain hypothesis [e,g., Litt et al., 2006; Tegmark, 2000], because the brain operates in a hot and uncontrolled environment and quantum superposition states needed for quantum computation would not last long enough to do any meaningful cognitive operations. More recently, new proposals have been put forward arguing that it is possible for the brain to maintain coherent superposition states long enough for quantum physical cognition to take place [Fisher, 2015; Weingarten et al., 2016].

Most researchers in quantum cognition do not rely on the hypothesis that the brain is some kind of quantum computer. Of course there are various opinions on the matter, and it is still an important open question, but quantum cognition research proceeds simply on the application of the mathematical principles to human and possibly other animal behavior. However, if quantum cognition researchers do not want to rely on the quantum brain hypothesis, how else would these computations be performed by the brain? Although this is still an important open question, several proposals have been made for describing how a classical neural system could possibly implement the computations required

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

for quantum probability theory [de Barros, 2012; de Barros and Suppes, 2009; Busemeyer et al., 2017; Khrennikov et al., 2018; Selesnick and Piccinini, 2018; Selesnick et al., 2017; Takahashi and Cheon, 2012]. Nevertheless Hameroff [2013] continues to argue that "From a well-known example of inductive reasoning: If the brain swims, looks, and quacks like a (quantum) duck, then it probably is a (quantum) duck" (p. 290). We will return to this controversial subject in Chapter 13.

1.3 Brief Overview

Although the later chapters of this book will go into much greater depth, it will be useful to first give a brief introduction to the fundamental aspects of quantum cognition, which comprise quantum probability theory, quantum dynamics, and quantum information processing.

1.3.1 What Is Quantum Probability?

What is quantum probability and how does it differ from classical probability? The purpose of a probability theory is to assign probabilities to events. For example, an event might be "the defendant committed the crime" and we may want to assign a probability (a number ranging between zero and one) to that event. A probability theory provides the rules (axioms) for making these assignments. Most social and cognitive scientists are only trained on classical probability, and so it comes as a surprise that other probability theories even exist. Like classical probability theory, quantum probability theory is based on a small set of axioms, but they happen to be different than the axioms of classical probability theory (also see Narens [2015] for a discussion of a variety of other probability theories). Chapter 2 presents more details about quantum probability theory, but here we present a few important ideas.

Classical probability has a long history, beginning in the seventeenth century with contributions by Pascal, Laplace, and other mathematicians, and it has continued to develop since that time. Much of classical probability theory was initially motivated by problems arising in classical physics, and later applications appeared in economics, engineering, statistics, cognitive science, and so on. The generally accepted axiomatic foundation for classical probability theory was proposed by Kolmogorov [1933/1950].

Classical probability theory is founded on the premise that events are represented as *subsets* of a larger set called the sample space. The collection of subsets forms a Boolean algebra that satisfies the axioms of closure (if A, B

are events, then $A \cap B$ is also an event), commutativity, $A \cap B = B \cap A$, and distributivity, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. A classical probability function is used to assign probabilities to events in an additive manner: If A and B are mutually exclusive, then the probability of A or B equals the probability of A plus the probability of B. Additivity is crucial for the theory to satisfy a standard test of "rationality" called the Dutch book test: A Dutch book (taking a gamble with a guaranteed loss) cannot be made against a person basing her decisions on expected value with an additive probability measure [de Finetti, 2017]. The classical law of total probability follows from the distributive property of events and the additive property of the probability function.

Quantum mechanics was invented by a brilliant group of physicists in the 1920s, including Planck, Einstein, de Broglie, Bohr, Heisenberg, Schrödinger, Born, Dirac, and others. Quantum theory was motivated by puzzling physical phenomena that ran counter to any kind of explanation in classical physics. Initially, quantum theorists were unclear about what they had created, but eventually it became clear that they had invented a new theory of probability. The generally accepted axiomatic foundation for quantum probability theory was proposed by Dirac [1930/1958] and von Neumann [1932/1955].

Quantum probability is founded on the premise that events are represented as *subspaces* of a vector space, called a Hilbert space (see Figure 1.1).[1] The collection of subspaces, however, does not form a Boolean algebra, and instead it forms a collection of overlapping Boolean algebras, called a partial Boolean algebra, which does *not* necessarily obey the closure, commutative, and distributive axioms. A quantum probability function is used to assign probabilities to events, which also satisfies additivity, and so quantum probability can also claim "rationality" on the basis of satisfying the Dutch Book test [Barnum et al., 2000; Pothos et al., 2017]. However, because the distributive axiom does not necessarily hold for quantum events, it follows that the law of total probability can be violated with quantum probabilities.

A famous theorem by Gleason [1957] tells us exactly how to compute the quantum probability function. We present more details in the next chapter, but Figure 1.1 illustrates the general idea using a three-dimensional space. Suppose a juror must decide between 'murder degree 1', 'murder degree 2', or 'innocent'. We can represent the event X = 'murder degree 1' by the ray (a one dimensional subspace) labeled X in the figure, and the event Y = 'murder degree 2' as the ray labeled Y in the figure, and the event Z = 'innocent' by the ray labeled Z in the figure. The event M = 'murder degree 1

[1] A Hilbert space is a vector space defined on a complex field endowed with an inner product. Completion of the inner product space is also required; however, quantum cognition researchers normally use finite-dimensional spaces, and the latter are always complete.

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

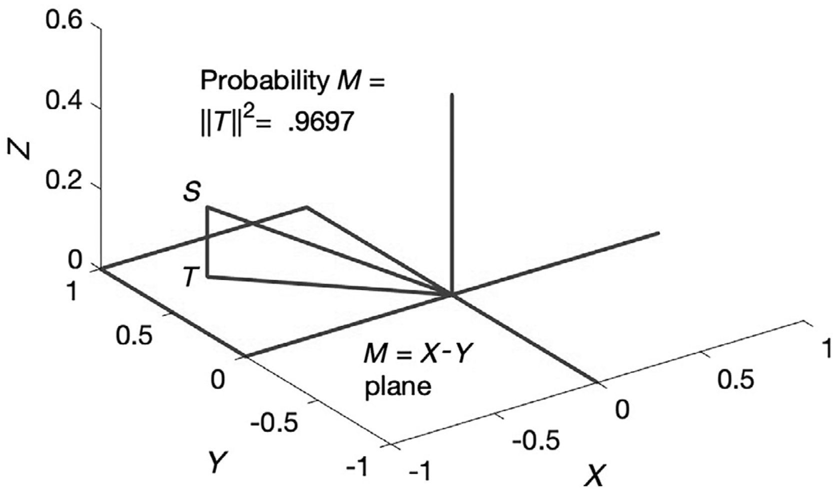

Figure 1.1 The rule for computing the probability of an event M given a state S. The event M is represented as a two-dimensional subspace (the X-Y plane) in a three-dimensional vector space. The state S is projected on the subspace M to produce the projection T, and the probability is the squared length of the projection $\|T\|^2$.

or murder degree 2' is then represented by the X-Y plane (a two-dimensional subspace) in the figure. The probabilities that a person assigns to these events are determined by a unit length state vector denoted S in the figure. The coordinates of a state with respect to some basis are called amplitudes. In general, the amplitudes can be complex numbers, although in this example they are real-valued. The probability of an event is obtained by (1) projecting the state on the subspace corresponding to an event, and then (2) squaring the length of the projection. For example, the probability of the event M is obtained by (1) projecting the state S onto the plane representing the event M producing the projection T, and then (2) computing the squared length of this projection, $\|T\|^2$. Gleason's theorem states that "in a separable Hilbert space of dimension at least three, whether real or complex, every measure on the closed subspaces is derived in this fashion" [see p. 885 in Gleason, 1957]. More informally, if we want to define events as subspaces instead of subsets, and we want the probabilities to satisfy additivity, then these probabilities must be computable from a single quantum state using a quantum algorithm as illustrated in Figure 1.1.[2]

[2] As discussed in Chapter 3, the quantum algorithm is more generally applied to probabilistic mixtures of superposition states represented efficiently by density matrices.

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

1.3.2 What Are Quantum Dynamics?

One can next ask: What are quantum dynamics and how do they differ from classical dynamics? The purpose of a probabilistic dynamical theory is to describe how probabilities assigned to states of a system change over time. Quantum dynamics are presented in Chapter 6, but here we present some basic ideas. Consider, for example, a juror who is evaluating evidence during a trial. Suppose the juror can express different degrees of belief in guilt on a rating scale ranging from 0 (definitely not guilty) to 100 (definitely guilty) in steps of 1 unit. While listening to the evidence, the juror's beliefs on this scale change over time. According to a classical probabilistic dynamic process, at each moment in time the juror is actually located at a specific level of belief, and the juror's belief moves from one belief level to another like a particle moving across time to produce a trajectory. In the left panel of Figure 1.2, the trajectory is the jagged path moving up and down and eventually drifting down toward not guilty. At the point in time when a decision is required, the juror simply reads out the existing level of belief to make a decision. The probabilities that are assigned to the belief levels across time for a classical dynamic system represent an outside observer's (e.g., a prosecutor's) uncertainty about a juror's trajectory.

According to a quantum dynamic process, at each moment in time the juror is superposed across belief levels. That is, at a particular moment in time

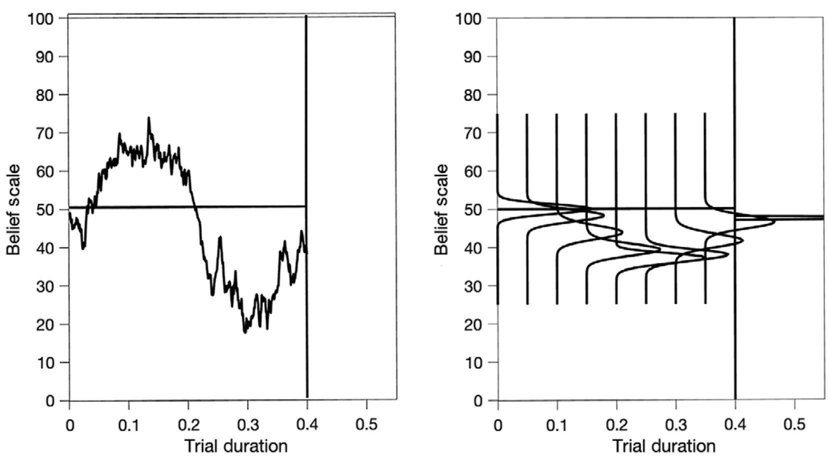

Figure 1.2 Markov random walk (left) and quantum walk (right) views of evidence accumulation. The vertical axis represents the different levels of beliefs and the horizontal axis represents time (the trial duration). The vertical line represents the time at which a decision must be made. In this example, the evidence is driving belief toward not guilty.

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

each belief level has an amplitude representing its potential to be selected. This superposition state (unit vector) is *rotated* (say from amplitudes favoring equally likely to amplitudes favoring innocent) across time to produce wave function. The right panel of Figure 1.2 illustrates the evolving probability distribution over belief levels produced by the rotating superposition state. At the point in time when a decision is required, the superposition state collapses to a definite decision. The amplitudes assigned to the belief levels across time represent the juror's uncertainty about her own beliefs.

Markov processes are very often used to model classical probabilistic dynamical systems. Such processes describe how a probability distribution across states evolves over time. Considering our juror example, the probability distribution represents an *external* observer's (e.g., the prosecutor's) uncertainty about where the juror's belief is located at a particular moment in time. A transition operator transforms the probabilities from one distribution at an earlier point in time to another distribution at a later point in time. The transition matrix is derived from a differential equation called the Kolmogorov forward equation.

Quantum processes describe how the juror's superposition state, an amplitude distribution, evolves over time. Considering our juror example again, the superposition state represents the juror's *internal* uncertainty over the belief scale at any moment. A unitary operator rotates the superposition state from one superposition state to another over time. The unitary operator is derived from a differential equation called the Schrödinger equation, named after the famous quantum theorist Erwin Schrödinger. (See Busemeyer et al. [2020a] for a comparison of Markov and quantum dynamic models.)

Recently a very general probabilistic dynamic model has been employed in cognitive research called an open systems model [Busemeyer et al., 2020a; Rivas and Huelga, 2012]. The name "open systems" comes from the idea that the cognitive system is influenced by both internal thought dynamics and external environmental influences. This provides a way to model both the external uncertainty about the state of the decision maker as well as the internal uncertainty that the decision maker has regarding his or her own beliefs. This more general model uses a differential equation called the Master equation that combines quantum and Markov processes by adding what are called Lindblad operators to the quantum process. Later in Chapter 7 we describe all of these dynamical models in more detail.

1.3.3 What Is Quantum Information Processing?

Finally, one can ask: What is quantum information processing and how does it differ from classical information processing? Information processing systems

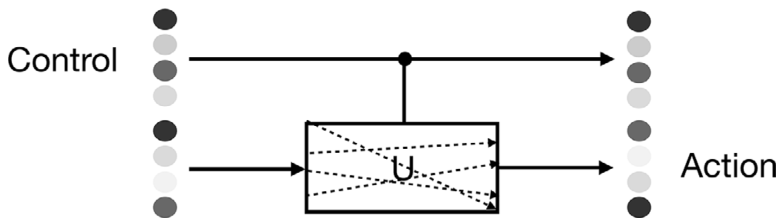

Figure 1.3 Controlled-U gate with a control input vector entering the upper line, and a target vector entering the U gate. The control input is passed unchanged to the output. The target vector may or may not be transformed by the U gate, depending on the control input vector.

generally accept input information, which is then used in a computation that eventually produces output actions. Information processing systems go beyond single-stage dynamics because they are used to perform multiple stages with sequences of actions in response to a sequence of inputs; for example, playing a game of chess. A classical example of an information processing system is a collection of if-then production rules like those used in classical cognitive architectures such as ACT-R [Anderson et al., 2004] or SOAR [Laird, 2012].

Quantum information processing operates according to a similar idea using controlled-U gates. A *controlled-U gate* is a unitary operator that takes a superposed vector as input and transforms it into another superposed output vector. The input vector is composed of two parts: The first part is the control signal, which determines whether or not to apply the U gate; the second part contains the target that is transformed by the U gate into an output vector, which is then used to determine the probability of actions. In Figure 1.3 the control input is a vector with coordinates represented by the upper column of shaded dots, and the target is another vector with coordinates represented by the lower column. The U gate in Figure 1.3 flips the target vector coordinates. Unlike symbolic production rules, but like connectionist networks, the inputs to a controlled-U gate are superpositions representing a distribution of uncertainty regarding the input features, and the outputs are also superpositions that produce probabilistic rather than deterministic actions. Also like connectionist networks, the unitary transformation matrix can be viewed as a network that connects inputs lines to output lines. Unlike connectionist networks, but like Bayesian reasoning, the probabilities computed from quantum information processing systems obey a coherent set of probability axioms. Sequences of situations and actions are generated by concatenating a sequence of controlled-U gates.

The field of quantum computing and quantum information is very advanced. Quantum computers are universal computing systems with the same computational power as classical computers, but with possibly faster speed of computations [Nielsen and Chuang, 2000]. The work in quantum cognition does not assume or require a quantum computer. Instead it uses these principles to build information processing systems designed to model situation–action sequences performed during problem solving and dynamic decision making by humans. Quantum information processing models applied to cognitive science are presented in Chapter 8.

1.4 Is There a Quantum Measurement Problem in Quantum Cognition?

As we discussed earlier, one of the key features of quantum theory is that measurement of a system changes the behavior of the system. There are at least two different ways to interpret the effect of measurement (see chapter 18 of Griffiths [2003] for a discussion). One common interpretation (by the way, not preferred by Griffiths for quantum physics) is that the observation causes a superposition state, containing the disposition for many states to be measured, to reduce to a definite state that will be observed in the actual measurement. The superposition state of the system before observation is sometimes viewed as a "quantum wave" and the transition from a superposition state to definite state associated with the observed outcome is often called the "collapse of the wave." This interpretation suggests that the superposition state has some type of actual physical existence (perhaps as an unstable neural state) and a physical collapse actually occurs when a measurement is made (pushing the neural dynamics into an attractor state). Another interpretation (preferred by Griffiths for quantum physics) is that a measurement determines the "preparation state," which is then used to compute the outcome probabilities, and an observed outcome is used to form a new conditional state. The latter interpretation suggests that the quantum state is a pre-probability used by a person to make predictions, and a measurement outcome is used to form a new conditional state for computing future predictions. Either one of these interpretations appears in various works on quantum cognition.

At this point, it is important to clarify what counts as a measurement in quantum cognition. Let's reconsider the experiments on the disjunction effect to see how this idea was applied. In the Shafir and Tversky [1992] experiment, the player is informed by the experimenter about the opposing player's move. In this case, the experimenter provides information about the opponent that is used to update the initial state of the player before the player takes an action in the prisoner's dilemma (PD) game. If we view the state as a superposition over the possible moves of the opponent, and we assume that the player actually

https://doi.org/10.1017/9781009205351.002 Published online by Cambridge University Press

accepts the experimenter's statement about the opponent's move, then the superposition should break down, aligning the initial state with the known move of the opponent. If the initial state is aligned with the known move of the opponent, then this has the same effect as a measurement that collapses the state on known move of the opponent player. Next consider the experiment by Croson [1999], which required predicting the move of the opponent. In this case, the experimenter requests a prediction and records a response from the player, and so it seems rather clear that an observation or measurement has taken place. By doing this, however, we are assuming that the superposition state over opponent moves is broken down and *revised* so that it aligns with prediction before taking the next action. In other words, it is assumed that if the player predicts the opponent will defect, then the player acts accordingly. It is as if the prediction is used as information to change the state. Finally, perhaps a person can ask him or herself a question (such as "What will my opponent do?") and thus perform a "self measurement." It is unclear whether or not this last example would result in a collapse or conditioning of the state. From these three examples, we see that the question of when a measurement is taken, or alternatively when a state is revised conditioned on new information, is a very important issue for quantum cognition.

1.5 Brief History and Chapter Summary

Quantum physical models of the brain have been discussed for over 50 years, perhaps beginning with [Ricciardi and Umezawa, 1967] with later work by [Jibu and Yasue, 1995] (also see Vitiello [2001] for a more accessible presentation of this theoretical work). However, as we mentioned earlier, models of quantum cognition only rely on the abstract mathematics of quantum theory, and are not directly based on quantum physical laws.

The idea of applying the abstract mathematics of quantum theory outside of physics to cognition and decision making began later, in the 1990s. Some of the earliest proposals were by Aerts and Aerts [1995] and Bordley [1998; Bordley and Kadane, 1999] on decision making; Roy and Kafatos [1999] and Conte et al. [2007] on perception; Gabora and Aerts [2002] to conceptual reasoning, and Khrennikov [1999] and Atmanspacher and Römer [2002] to psychology more generally.

The field began growing with a series of international meetings and their proceedings [see Bruza et al., 2007, for the first] held each year, called the Quantum Interaction Conference. A special issue appeared in 2009 in the *Journal of Mathematical Psychology* [Bruza et al., 2009] that provided a solid foothold for the field to move forward, and numerous special issues have appeared on the topic since that time [see, e.g., Gray, 2013]. General reviews of the field started appearing beginning with Pothos and Busemeyer [2013], later Ashtiani and Azgomi [2015] and Bruza et al. [2015a], and most recently Pothos and Busemeyer [2022]. Day-long in-person tutorials have been held almost every year at the Cognitive Science Society beginning in 2007, and published tutorials have appeared in Yearsley and Busemeyer [2016] and later Yearsley [2017]. Books on quantum cognition and decision have been written including the first edition of this book [Busemeyer and Bruza, 2012] as well as an earlier book by Khrennikov [2010] and a later one by Bagarello [2019]. Van Rijsbergen [2004] introduced the application of quantum principles to information retrieval. A related area to quantum cognition known as "quantum social science" has also emerged [Haven and Khrennikov, 2013; Wendt, 2015].

In summary, we began this chapter by discussing reasons for considering a quantum program of research for cognition. Motivating psychological reasons included applying the concepts of superposition to represent decision-maker states of uncertainty, sensitivity to measurement that produces interference effects, and the correspondence principle that is closely related to incompatible measurements and sequential effects of measurement. We also discussed reasons based on the contextual nature of judgments and decisions and pointed out how the concept of entanglement can be used to account for these effects. Two other important issues were addressed: Why evolution might generate a quantum reasoning system when we interact with a classical world, and how the brain might perform the computations required for quantum reasoning. The chapter also provided a overview of the three main topics within quantum cognition by briefly reviewing the basic ideas behind quantum probability, quantum dynamics, and quantum information processing. In addition, we examined how these three quantum principles compare to classical probability, dynamics, and information processing. The chapter included an example application of quantum cognition modeling to a puzzling phenomenon in decision making called the disjunction effect, and we discussed some of the measurement issues that arose with this application. Hopefully this introduction has generated sufficient interest and momentum for the reader to continue and discover the broad variety of applications of quantum theory to human cognition.