

Quantum dynamics of human decision-making

Jerome R. Busemeyer*, Zheng Wang, James T. Townsend

Indiana University, USA

Received 10 December 2004; received in revised form 11 January 2006

Available online 15 March 2006

Abstract

A quantum dynamic model of decision-making is presented, and it is compared with a previously established Markov model. Both the quantum and the Markov models are formulated as random walk decision processes, but the probabilistic principles differ between the two approaches. Quantum dynamics describe the evolution of complex valued probability amplitudes over time, whereas Markov models describe the evolution of real valued probabilities over time. Quantum dynamics generate interference effects, which are not possible with Markov models. An interference effect occurs when the probability of the union of two possible paths is smaller than each individual path alone. The choice probabilities and distribution of choice response time for the quantum model are derived, and the predictions are contrasted with the Markov model.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Quantum; Markov; Dynamics; Decision-making; Random-walk; Diffusion; Choice; Response-time

0. Introduction

Many of the dynamical models being used in psychology today had their beginnings in theoretical physics. In particular, Markov processes and diffusion models, currently used to model psychological phenomena, were originally formulated to explain physical phenomena. For example, the Markov model used by Falmagne to model preference evolution (Falmagne, Regenwetter, & Grofman, 1997; Regenwetter, Falmagne, & Grofman, 1999) was initially developed as a kinetic theory of gases (by Poisson, Gibbs, and Boltzman); and the diffusion model which is currently used to model signal detection (see Ratcliff & Smith, 2004, for a review) was originally developed to explain Brownian motion (by Einstein, Ornstein, Uhlenbeck, Wiener).

Markov processes and diffusion models are based on a ‘single path’ assumption that lies at the heart of the classical mechanical theory of physics. During a single realization or sample, the Markov process can be in one and only one basis state at any moment in time, and the

process jumps from one basis state to another across time. For example, if the basis states represent different levels of evidence for a hypothesis, then the Markov process produces a path consisting of a sequence of evidence states.

Perhaps it is best to describe this assumption by way of a simple physical example known as the paradox of the recombined beams (see French & Taylor, 1978, Chapter 7, p. 284). Referring to Fig. 1, a single plane polarized photon is dispersed from a light source in state X_0 , from which it can reach either of two separate channels rotated with respect to the source. Finally, the photon can pass through either channel to a detector, orientated in an opposite direction relative to the source, from which the photon may be detected (state X_4) or not (state X_3). If one observes inside the box to see which channel the photon enters, then there is a probability p that it passes through the first channel, and there is a probability $(1-p)$ that it passes through the other channel (because of the rotation of the filters), but one never sees it pass through both. If the photon is observed to pass through the first channel, then the probability of eventually reaching the detector is $(1-p)$; if the photon is observed to pass through the second channel, then the probability of eventually reaching the detector is p .

Now suppose that no measurement is made to determine which channel the photon passes through. According to the

*Corresponding author. Psychology Department, Indiana University, Bloomington, IN 47405, USA. Fax: +812 855 4691.

E-mail address: jbusemey@indiana.edu (J.R. Busemeyer).

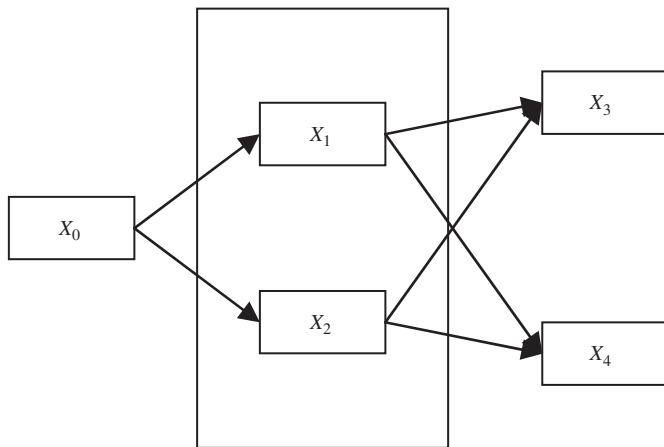


Fig. 1. Illustration of various possible sample paths for a dynamic system.

classic theory, the photon could either travel through the $X_0 \rightarrow X_1 \rightarrow X_4$ path with probability $p \cdot (1-p)$, or it could travel through the $X_0 \rightarrow X_2 \rightarrow X_4$ path with probability $(1-p) \cdot p$, but it cannot travel through both. Thus, according to the classic theory, the probability of the union of these two mutually exclusive events is the sum, $2 \cdot p \cdot (1-p)$. In fact, however, no photons are ever observed to reach the detector under this condition!

Modern dynamic models of physics are currently based on the theory of *quantum* mechanics. This modern theory rejects the single path assumption. Referring to Fig. 1, if we do not measure the events inside the box, then we cannot say the state of the system must either transit from X_0 to state X_1 exclusively or from X_0 to state X_2 exclusively to reach state X_4 . If we do not observe what happens inside the box of Fig. 1, then the system enters into a “superposition” of states X_1 and X_2 , which is not equal to either one. The amazing consequence of this new assumption is that the probability of going from X_0 to X_4 by either of the two hidden paths can be *less* than the probability of a single path. This is called the interference property of quantum mechanics (see Feynman, Leighton, & Sands, 1966, Chapters 1, 3, 4 and 5 for many examples).

This fundamental difference in assumptions between the classic versus the modern view of dynamics leads to the main question to be addressed in this article: Is it more suitable to model brain processes using quantum waves or Markov particles (see Glimcher, 2005; Penrose, 1989; Pribram, 1991)? Psychologists have already applied the Markov approach with considerable success, but they have rarely examined the quantum approach. Can we derive a meaningful model of human information processing from quantum dynamic principles? And if this is possible, how does it differ qualitatively from the models derived from the classical dynamic principles?

First we review and compare the basic principles of Markov versus quantum dynamic theories and note their similarities and dissimilarities. Then we illustrate the application of quantum dynamics by developing a quantum random walk model of decision-making and compare

this with a previously established Markov random walk decision model. Finally, we review several other potentially useful applications of the quantum dynamic approach to human information processing.

1. Comparison of quantum versus Markov dynamics

1.1. Basic principles

Markov principles. Markov dynamics are derived from three simple but basic rules (see Bhattacharya & Waymire, 1990; Cox & Miller, 1965; Kemeny & Snell, 1960; Kolmogorov, 1933). Suppose $|\psi\rangle$, $|\chi\rangle$, and $|\phi\rangle$ are mutually exclusive states of a Markov state space.¹

1. The probability of the transition $|\psi\rangle \rightarrow |\chi\rangle$ equals the conditional probability $\Pr[\chi|\psi]$. This is a positive real number between zero and one.
2. The probability of $|\psi\rangle \rightarrow |\chi\rangle$ and $|\chi\rangle \rightarrow |\phi\rangle$ is equal to $\Pr[\phi|\chi] \cdot \Pr[\chi|\psi]$.
3. The probability of $|\psi\rangle \rightarrow |\phi_2\rangle \rightarrow |\chi\rangle$ or $|\psi\rangle \rightarrow |\phi_1\rangle \rightarrow |\chi\rangle$ is equal to $\Pr[\chi|\phi_2] \cdot \Pr[\phi_2|\psi] + \Pr[\chi|\phi_1] \cdot \Pr[\phi_1|\psi]$.

For example, apply these rules to the problem shown in Fig. 1, $\Pr[X_1|X_0] = p$, $\Pr[X_4|X_1] = (1-p)$, $\Pr[X_2|X_0] = (1-p)$, and $\Pr[X_4|X_2] = p$. The probability of reaching the detector through either of the two paths is $\Pr[X_0 \rightarrow X_1 \rightarrow X_4 \cup X_0 \rightarrow X_2 \rightarrow X_4] = \Pr[X_0 \rightarrow X_1 \rightarrow X_4] + \Pr[X_0 \rightarrow X_2 \rightarrow X_4] = \Pr[X_4|X_1] \cdot \Pr[X_1|X_0] + \Pr[X_4|X_2] \cdot \Pr[X_2|X_0] = 2 \cdot p \cdot (1-p)$ which of course must exceed $\Pr[X_0 \rightarrow X_1 \rightarrow X_4] = \Pr[X_4|X_1] \cdot \Pr[X_1|X_0] = p \cdot (1-p)$.

Quantum principles. Quantum dynamics are derived from a slightly different set of basic rules (see Dirac, 1958; Feynman et al., 1966; French & Taylor, 1978; Nielsen & Chuang, 2000; Shankar, 1994; Von Neumann, 1955). According to quantum theory, *unobserved* state transitions obey probability amplitude rules. The probability of an observed event is equal to the squared magnitude of the corresponding probability amplitude of the event. Suppose $|\psi\rangle$, $|\chi\rangle$, and $|\phi\rangle$ are distinct quantum states (unit length complex vectors).

1. The probability amplitude of making a transition $|\psi\rangle \rightarrow |\chi\rangle$ is equal to the inner product $\langle \chi|\psi\rangle$. This is a complex number with magnitude less than or equal to one.
2. The probability amplitude of $|\psi\rangle \rightarrow |\chi\rangle$ and $|\chi\rangle \rightarrow |\phi\rangle$ is equal to $\langle \phi|\chi\rangle \cdot \langle \chi|\psi\rangle$.
3. The probability amplitude of $|\psi\rangle \rightarrow |\phi_2\rangle \rightarrow |\chi\rangle$ or $|\psi\rangle \rightarrow |\phi_1\rangle \rightarrow |\chi\rangle$ is equal to $\langle \chi|\phi_2\rangle \cdot \langle \phi_2|\psi\rangle + \langle \chi|\phi_1\rangle \cdot \langle \phi_1|\psi\rangle$.

¹The ket $|\psi\rangle$ is related to the concept of a column vector, and the bra $\langle \chi|$ is related to the concept of a row vector, and the inner product $\langle \chi|\psi\rangle$ is a number. See the appendix for more details.

Note these rules are very similar to the rules for Markov processes except for the fact that probability amplitudes are complex numbers rather than positive real numbers. Both have magnitudes less than or equal to unity.

As an example, let us apply these rules to problem shown in Fig. 1. Consider first the case where we *observe* which path the photon travels. In this case, we observe one path with probability $|\langle X_4|X_1\rangle\langle X_1|X_0\rangle|^2 = |\langle X_4|X_1\rangle|^2|\langle X_1|X_0\rangle|^2 = p \cdot (1-p)$, or we observe the other path with probability $|\langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2 = |\langle X_4|X_2\rangle|^2|\langle X_2|X_0\rangle|^2 = (1-p) \cdot p$. Thus, when we observe the path that is traveled, the total probability of starting at X_0 and ending at X_4 is the sum of the two observed probabilities, $|\langle X_4|X_1\rangle\langle X_1|X_0\rangle|^2 + |\langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2 = 2 \cdot p \cdot (1-p)$, which agrees with Markov theory. Next consider the case where we *do not observe* which path the photon travels. In this case, the probability of starting at X_0 and ending at X_4 is obtained first by summing the two probability amplitudes, and then taking the square magnitude, $|\langle X_4|X_1\rangle\langle X_1|X_0\rangle + \langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2$, which disagrees with Markov theory. How can the latter probability for two unobserved paths be smaller than the probability of a single path? Expansion yields

$$\begin{aligned} &|\langle X_4|X_1\rangle\langle X_1|X_0\rangle + \langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2 \\ &= |\langle X_4|X_1\rangle\langle X_1|X_0\rangle|^2 + |\langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2 \\ &\quad + 2 \cdot |\langle X_4|X_1\rangle\langle X_1|X_0\rangle| \cdot |\langle X_4|X_2\rangle\langle X_2|X_0\rangle| \cdot \cos(\omega), \end{aligned}$$

where ω is the angle between $\langle X_4|X_1\rangle\langle X_1|X_0\rangle$ and $\langle X_4|X_2\rangle\langle X_2|X_0\rangle$ in the complex plane (see Appendix A). The third term in this expansion, called the interference term, can be positive or negative, and this makes quantum theory differ from Markov theory. In particular, if $\omega = 180^\circ$, then $\cos(\omega) = -1$ and the third term cancels the first two terms so that the probability is reduced to zero. Interference effects only occur when we do not observe the path that the system travels.²

1.2. State representation

Both the Markov and the quantum dynamic models postulate a set of basis states. The basis states are used to represent the various mental states that, at least hypothetically, can be measured or reported to occur during an experiment. Many different psychological applications can be envisioned in terms of basis states. For a signal detection task, the basis states represent different levels of evidence strength favoring one hypothesis over another (Laming, 1968; Link & Heath, 1975; Ratcliff, 1978; Smith, 2000). For a binary choice task, the basis states represent different levels of preference strength for one option over another (Bockenholt, 2001; Busemeyer & Townsend, 1993; Diederich, 1997; Townsend & Busemeyer, 1995). For a voting task, the basis states represent different partial orders of

²When the paths are unobserved, the system enters a superposition state: $|X_0\rangle = u_{10}|X_1\rangle + u_{20}|X_2\rangle = u_{10} \cdot (u_{31}|X_3\rangle + u_{41}|X_4\rangle) + u_{20} \cdot (u_{32}|X_3\rangle + u_{42}|X_4\rangle) = (u_{31} \cdot u_{10} + u_{32} \cdot u_{20})|X_3\rangle + (u_{41} \cdot u_{10} + u_{42} \cdot u_{20})|X_4\rangle$.

preferences over candidates (Falmagne et al., 1997; Regenwetter et al., 1999). For an information processing task, the basis states represent different combinations of completed and uncompleted tasks (Batchelder & Reiffer, 1999; Fisher & Goldstein, 1983).

To build connections between the Markov and quantum systems, we will assume that states are abstract elements of an m dimensional vector space.³ We form a basis for this vector space by a choosing a set of m linearly independent basis vectors:

$$\Omega = \{|1\rangle, |2\rangle, \dots, |j\rangle, \dots, |m-1\rangle, |m\rangle\}.$$

However, any linear independent set can be transformed into an orthonormal set, and so we will assume that the basis vectors are orthonormal. The basis vector $|j\rangle$ corresponds to the j th possible measurement outcome: If the individual is in basis state $|j\rangle$, and we immediately take a measurement, then we will certainly observe the j th measurement outcome.⁴

Markov states. The pure state of the Markov system at time t can be defined as a convex combination of the basis states in Ω :

$$|P(t)\rangle = \sum_{j \in \Omega} w_j(t) \cdot |j\rangle,$$

where $w_j(t)$ is a random indicator variable with values equal to 1 or 0 depending on whether or not the individual is in state $|j\rangle$ at time t . According to the Markov model, the individual must be in one and only one state at each moment in time so that $\sum w_j(t) = 1$. The probability of being in state $|j\rangle$ at time t is denoted $P_j(t) = \text{Pr}[w_j(t) = 1]$.

The *mixed* state of a Markov system is defined by the column vector, $P(t)$, where the probability $P_j(t)$ is the coordinate in the j th row corresponding to the basis vector $|j\rangle$. The probabilities in $P(t)$ sum to unity, $\sum P_j(t) = 1$, so that $P(t)$ represents the probability distribution across the basis states at time t . The set of all probability distributions defines the (mixed) state space of the Markov system.

Quantum states. A *superposition* state of the quantum system at any moment in time t is defined as a linear combination of the basis vectors in Ω :

$$|\psi(t)\rangle = \sum \langle j|\psi(t)\rangle \cdot |j\rangle = \sum \psi_j(t) \cdot |j\rangle,$$

where $\psi_j = \langle j|\psi\rangle$ is the probability amplitude of observing the basis state $|j\rangle$ at time t . The probability amplitude, ψ_j , is a complex number with a magnitude less than or equal to one. The column vector ψ represents the quantum state using coordinates defined by the Ω basis, where ψ_j is the

³See Appendix A for a summary of linear algebra on vector spaces. For simplicity, we initially limit the discussion to a finite number of basis states. Note that m can be quite large to approximate a continuum. Later in Section 1.6 we will consider a continuum of states.

⁴In Quantum theory, An *observable* is defined by the Hermitian linear operator $\mathbf{M} = \sum_j f(j) \cdot |j\rangle\langle j|$, where f is a real valued function of the state index, and $|j\rangle\langle j|$ is a projection operator. $\mathbf{M}|j\rangle = f(j) \cdot |j\rangle$ projects each basis state onto itself, and each measurement outcome is recovered from the inner product $\langle j|\mathbf{M}|j\rangle = f(j)$.

coordinate in the j th row corresponding to the basis vector $|j\rangle$. The squared length of ψ must be unity, $|\psi|^2 = 1$, so that the collection of squared amplitudes produces a probability distribution over the basis states. The state space of the quantum system is defined as the set of all linear combinations of the basis vectors with unit length.

State interpretations. The superposition state ψ of the quantum model corresponds to the mixed state P of the Markov model. However, these representations of states are mathematically very different: ψ is a complex vector of unit length, whereas P is a non-negative real vector that sums to unity. Conceptually, these states are also radically different. According to the Markov model, for any given realization, the unobserved system occupies exactly one basis state $|j\rangle$ at each moment in time. A sample path of the Markov process is a series of jumps from one basis state to another, which moves like a bouncing particle across time. A different path is randomly sampled for each realization of the Markov process. According to the quantum model, for any given realization, the unobserved system does not occupy any particular basis state at each moment in time. A realization of the quantum process is a fuzzy spread of membership across the basis states, which moves like a traveling wave across time. Each realization of the quantum process is identical, producing the same series of states across time. All of the randomness in the quantum model results from taking a measurement.

1.3. State transitions

Markov transitions. A mixed Markov state can be transformed into another mixed state by a transition matrix, symbolized as T . Each element of the transition matrix, $T_{ij}(t)$, defines the transition probability to state $|i\rangle$ from state $|j\rangle$ during time t , for every $(i, j) \in \Omega^2$. In other words, $T_{ij}(t) = \text{Pr}[\text{state } |i\rangle \text{ at time } t \text{ given state } |j\rangle \text{ at time } 0]$. Starting from state $|j\rangle$ the system must land into one of the possible basis states in Ω at time t , and therefore the columns of the transition matrix must sum to one.⁵

According to the basic principles, the total probability of ending at state $|i\rangle$ at time t is obtained from the sum across all the possible starting positions

$$P_i(t) = \sum T_{ij}(t) \cdot P_j(0).$$

Thus the probability distribution at time t is related to the initial distribution by the linear equation

$$P(t) = T(t) \cdot P(0). \tag{1}$$

Quantum transitions. A quantum state can be transformed into another state by a unitary operator, symbolized as U , with $U^\dagger U = I$, which is required to preserve

⁵Note that this definition of the transition matrix (transiting from column to row) is the transpose of the definition normally used in stochastic processes texts. This was chosen, however, to match the definitions typically used in quantum mechanics texts.

inner products: If we transform $|\chi\rangle$ and $|\psi\rangle$ to $U|\chi\rangle$ and $U|\psi\rangle$, then $\langle\chi|U U|\psi\rangle = \langle\chi|\psi\rangle$. In particular, unitary operators preserve lengths: $\langle\psi|U U|\psi\rangle = \langle\psi|\psi\rangle = 1$.

The probability amplitude of starting in state $|\psi\rangle$ and passing through U and then observing basis state $|i\rangle$ can be expressed in terms of the Ω basis as

$$\langle i|U|\psi\rangle = \sum_{j \in \Omega} \langle i|U|j\rangle \langle j|\psi\rangle = \sum_{j \in \Omega} u_{ij} \cdot \langle j|\psi\rangle.$$

This follows directly from the basic principles. The transition probability amplitude, $u_{ij} = \langle i|U|j\rangle$, represents the probability amplitude of transiting to basis state $|i\rangle$ from basis state $|j\rangle$ going through the unitary operator U . Thus the right-hand side is the sum of all the path probability amplitudes from $|\psi\rangle$ through $|j\rangle$ to $|i\rangle$, and each path probability amplitude is the product of the link probability amplitudes, $\langle i|U|j\rangle \langle j|\psi\rangle$.

The above equation implies that the unitary operator is a linear operator. The transition probability amplitudes form a matrix U , with u_{ij} in row i column j , representing the coordinates of U with respect to the Ω basis, so that we can express the effect of the unitary operator as the linear transformation: $\chi = U \cdot \psi$.

If the state $|\psi(0)\rangle$ is processed by the unitary operator U for some period of time t , then it produces the new state $|\psi(t)\rangle = U(t)|\psi(0)\rangle$. Using the coordinates defined with respect to the Ω basis, the state transition rule is expressed by the following linear transition equation:

$$\psi(t) = U(t) \cdot \psi(0). \tag{2}$$

Note that the transition rule for probability amplitudes in quantum processes (Eq. (2)) corresponds to the transition rule for probabilities in Markov processes (Eq. (1)). However the former model operates on complex amplitudes whereas the latter model is restricted to positive real numbers. In terms of this property, the quantum model is more general than the Markov model.

1.4. Dynamical equations

Markov dynamics. From basic principles of Markov processes, it follows that the transition probabilities satisfy an important property known as the Chapman–Kolmogorov equation

$$T_{ij}(s+t) = \sum_{k \in \Omega} T_{ik}(s) \cdot T_{kj}(t)$$

or in matrix form

$$T(s+t) = T(s) \cdot T(t). \tag{3}$$

This is known as the semigroup property of dynamic systems, and from this semigroup property one can derive a differential equation known as the Kolmogorov forward equation (Bhattacharya & Waymire, 1990, p. 266):

$$\frac{d}{dt} T(t) = Q \cdot T(t), \tag{4a}$$

Multiplying both sides of Eq. (4a) by $P(0)$ and setting $P(t) = T(t) \cdot P(0)$ yields

$$\frac{d}{dt} P(t) = Q \cdot P(t). \quad (4b)$$

The *intensity* matrix⁶ Q has elements $q_{ij} = \lim_{\tau \rightarrow 0} (T_{ij}(\tau) - T_{ij}(0))/\tau$ in row i column j which represents the instantaneous rate of change to $|i\rangle$ from $|j\rangle$. These rates control the flow of probability over time.

Quantum dynamics. If the state $|\psi(0)\rangle$ is processed by the unitary operator U for some period of time t , and it is immediately processed again by the same unitary operator U for an additional time s , then it produces the new state $|\psi(s+t)\rangle = U(s+t)|\psi(0)\rangle = U(s)U(t)|\psi(0)\rangle$. Using the matrix coordinates of the operator, this implies

$$u_{ij}(s+t) = \sum_k u_{ik}(s) \cdot u_{kj}(t)$$

or more generally,

$$U(s+t) = U(s)U(t). \quad (5)$$

This follows directly from the basic principles of quantum dynamics. Eq. (5) for the quantum dynamic model corresponds to Eq. (3) of the Markov process model. In other words, this is the quantum analogue of the Chapman–Kolmogorov equation: the unitary operator for the quantum dynamic model satisfies the same semigroup property as does the transition matrix for the Markov process model (see Gardiner, 1991, p. 165; Hughes, 1989, p. 114).

It follows that the unitary operator satisfies the following differential equation known as the Schrödinger equation:

$$\frac{d}{dt} U(t) = -i \cdot H \cdot U(t). \quad (6a)$$

Multiplying both sides of Eq. (6a) by $\psi(0)$ and setting $\psi(t) = U(t) \cdot \psi(0)$ yields

$$\frac{d}{dt} \psi(t) = -i \cdot H \cdot \psi(t). \quad (6b)$$

The *Hamiltonian* H has elements $h_{ij} = \lim_{\tau \rightarrow 0} (U_{ij}(\tau) - U_{ij}(0))/\tau$ in row i column j representing the instantaneous rate of change to $|i\rangle$ from $|j\rangle$. Note that the Schrödinger equation (Eq. (6)) corresponds directly to the Kolmogorov forward equation (Eq. (4)). (The inclusion of the factor, $-i$, is needed to satisfy the unitary property.)

1.5. Evolution of probability distributions

For the Markov process, the solution to the forward equation is the matrix exponential

$$T(t) = e^{Q \cdot t}. \quad (7a)$$

Thus the solution for the probability distribution over time is directly obtained from

$$P(t) = e^{Q \cdot t} \cdot P(0). \quad (7b)$$

The solution to the Schrödinger equation is given by the matrix exponential of the Hamiltonian matrix:

$$U(t) = e^{-i \cdot H \cdot t}. \quad (8a)$$

Thus the probability amplitudes evolve across time according to the following equation:

$$\psi(t) = e^{-i \cdot H \cdot t} \cdot \psi(0). \quad (8b)$$

Note that the quantum solution given by Eq. (8b) corresponds directly to the Markov solution given in Eq. (7b). However, according to the quantum model, the final probability for each state is obtained by taking the squared magnitude of the corresponding probability amplitude: $P_j(t) = |\psi_j|^2$. Thus the probability distribution for the Quantum model is a nonlinear transformation of Eq. (8b).

1.6. Model parameters

Intensity matrix. The specification of the intensity rates in Q is critical for Markov models. The intensity matrix Q must satisfy the following constraints:

$$q_{ij} \geq 0 \text{ for } i \neq j \quad \text{and} \quad \sum_{i \in \Omega} q_{ij} = 0,$$

where the first inequality is required to make the transition probabilities non-negative, and the latter is required because the transition probabilities within a column sum to one. Note that the intensity matrix Q of the Markov model is allowed to be asymmetric.

Metric information about the distances between Markov states is represented by the intensity matrix. In the case of a random walk type model, transitions only occur to adjacent states. In this case, the intensity matrix Q is a systolic matrix: $q_{ij} = 0$ for states $|i\rangle$ and $|j\rangle$ that are more than one step apart.

Hamiltonian matrix. The specification of the rates in H is also critical for the quantum model. The Hamiltonian matrix H must be Hermitian, $H = H^\dagger$, to guarantee that the solution of Eq. (8a) is a unitary matrix. Note that the Hamiltonian matrix H must obey different constraints than the intensity matrix Q , and therefore, the Markov model is not a special case of the quantum model.

Metric information about the distances between quantum states is represented by the Hamiltonian matrix. Analogous to the Markov model, for a random walk type process, we assume that probability amplitudes only diffuse to adjacent states. Thus, the Hamiltonian is a systolic matrix: $h_{ij} = 0$ for states $|i\rangle$ and $|j\rangle$ that are more than one step apart. (This corresponds to the Hamiltonian used in the crystal model discussed by Feynman et al., 1966, Chapter 16.)

1.7. Continuous state models

It is straightforward to extend the discrete state random walk Markov model to allow for a continuum of states.

⁶Also called the infinitesimal generator of the transition matrix.

If we index the basis states as $x = j \cdot \Delta$ and we let the step size $\Delta \rightarrow 0$ in the limit, then the discrete state process converges in probability distribution to a continuous state process (see Bhattacharya & Waymire, 1990, pp. 386–388). For example, using the intensity matrix defined in Section 2.1 below, the Kolmogorov forward equation converges to the following partial differential equation (see Bhattacharya & Waymire, 1990, pp. 386–388; also see Appendix D):

$$\frac{\partial}{\partial t} P(t, x) = \frac{\sigma^2}{2} \cdot \frac{\partial^2}{\partial x^2} P(t, x) - \mu \cdot \frac{\partial}{\partial x} P(t, x). \tag{9}$$

From this equation, it is possible to derive the first passage time distribution for the diffusion model used by Ratcliff (1978).

The quantum random walk model can also be extended to a continuum of states. Once again we define the level of confidence as $x = j \cdot \Delta$ and we let the step size $\Delta \rightarrow 0$ in the limit, in which case the discrete state process converges to a continuous state process (see Feynman et al., 1966, Chapter 16). Using the Hamiltonian defined in Section 2.2 below, the Schrödinger equation converges to the following partial differential equation (see Appendix D)

$$-i \frac{\partial}{\partial t} \psi(t, x) = \sigma^2 \frac{\partial^2}{\partial x^2} \psi(t, x) - v(x) \cdot \psi(t, x). \tag{10}$$

Again there is a striking similarity between the forms of the two models, except that P is a real valued probability density, and ψ is a complex valued probability amplitude density. The main difference occurs with the terms associated with μ in Eq. (9) and $v(x)$, in Eq. (10). The latter difference arises from the fact that the Markov process has an asymmetric intensity matrix, whereas the quantum process has variation along the diagonal of the Hamiltonian. Thus, the essential ideas presented above for the discrete state models generalize to the continuous state models.

1.8. Summary

If one already knows the basic principles of Markov dynamics, then it is fairly simple easy to understand the basic principles of quantum dynamics. Essentially quantum dynamics follow the same rules as Markov dynamics except that the rules are applied to probability amplitudes rather than probabilities, and probabilities are obtained by squaring the amplitudes. It is possible to form a hybrid model that combines the Markov and quantum processes into a unified probabilistic dynamic system. Essentially, this is created by forming a probability mixture over quantum states. A density operator is used in quantum theory for this purpose (Gardiner, 1991; Von Neumann, 1955).

2. Random walk signal detection model

Now we develop a more specific application of the Markov and quantum random walk models to the problem of signal detection. In this decision situation, a human operator monitors information for potential targets (mis-

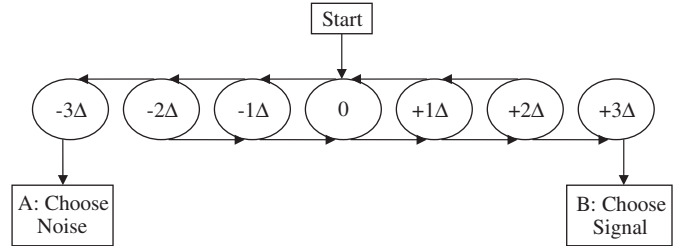


Fig. 2. Random walk signal detection decision model. This example has only 7 states. The calculations reported later use 51 states.

siles in the military, or tumors in medicine). Incoming information provides uncertain and conflicting evidence, but at some point in time, a decision must be made, and the sooner the better. Incorrectly deciding that the target is an enemy (i.e., a false alarm) could result in deadly consequences, but so could the error of misidentifying a real foe as a friendly agent (i.e., a miss). The goal of a psychological theory about this process is to describe the probability of making each choice, the confidence ratings assigned to these decisions, as well as the distribution of time that it takes to make the decision.

Basis states. The basis states are selected to represent all the possible measurement outcomes produced by the choice and confidence rating measures. A binary choice (signal versus noise) is represented by two orthonormal basis states, $|+\rangle$ for signal and $|-\rangle$ for noise, which forms a two dimensional vector space. The levels of confidence are represented by m orthonormal basis states $|1\rangle, |2\rangle, |3\rangle, \dots, |m\rangle$, which forms a m -dimensional vector space. The tensor product of these two vector spaces forms a $2m$ dimensional vector space. Also, to allow for a special neutral preference state with zero confidence in either option, we append a neutral state $|0\rangle$ to the basis set. Thus the complete set of basis states is defined as

$$\Omega_{SD} = \{|-m\rangle, |-m+1\rangle, \dots, |-1\rangle, |0\rangle, |+1\rangle, \dots, |+m-1\rangle, |+m\rangle\}.$$

Fig. 2 is a state transition diagram that illustrates the states and transitions between states for a random walk decision process. This figure shows only seven states, but m may be much larger to approximate a continuum. All of the computations reported below are actually based on $m \geq 25$ confidence levels.

The use of a common basis vector to simultaneously represent a particular choice and a particular confidence level has important empirical implications. According to quantum theory, we are assuming that the choice and confidence are *compatible* measurements. That is, we assume that we can access choice and confidence in parallel. The measurement of one does not affect the measurement of the other, and the order does not matter. Incompatible measures must be serially accessed in some particular order—one measurement disturbs the other, and the order of measurement is important. As we will see

later in this section, this assumption makes predictions concerning the effect or lack of effect of order that measurements are taken. This assumption may be wrong, which would indicate the need for a different type of representation.

2.1. Markov model

First we present the Markov version of the random walk signal detection model. Decades of theoretical work on this problem have shown that the Markov random walk model provides an impressive account of the choice probabilities and response time distributions observed with signal detection tasks (see Ratcliff & Smith, 2004, for a recent review).

At the beginning of the choice problem, the decision maker starts with an initial evidence level, labeled $X(0)$. For example, the decision maker may start at the neutral level, zero, in Fig. 2. Samples of evidence are presented to the decision maker over time (e.g., blips appearing occasionally on a radar screen). At some point in time, τ_1 , the first piece of evidence is sampled and observed, causing the initial level $X(0)$ to be incremented up or down producing a new level $X(\tau_1) = X(0) + \Delta(\tau_1)$. If the evidence favors one choice (e.g., option *B*, respond signal, decide it is an enemy agent), then the increment is positive ($\Delta(t) = +\Delta > 0$); if the evidence favors the other choice (e.g., option *A*, respond noise, decide it is a friendly agent), then it is negative ($\Delta(t) = -\Delta < 0$). Referring to Fig. 2, the level of evidence can move up or down a step—for example, the first piece of evidence may favor signal causing a step up from level 0 to level $+1\Delta$. Suppose after a period of time t , a total of n samples have been observed, and the evidence level is at $X(t)$; if the next piece of evidence is observed after period of time τ_{n+1} , then the new level becomes $X(t + \tau_{n+1}) = X(t) + \Delta(t + \tau_{n+1})$.

Initial state. The initial state existing at the beginning of the decision is the $2m+1$ vector, $P(0)$, representing the probability distribution across the basis states in Ω_{SD} . For many experiments, the decision maker starts without information and no prior knowledge, which is usually represented by a process that starts near the neutral state $|0\rangle$. In particular, for the applications presented below, we set $P_{-1}(0) = P_0(0) = P_{+1}(0) = 1/3$.

The effect of the initial distribution on the final distribution is easy to characterize in Markov models. Define $P(t; j)$ as the probability distribution across states at time t , given that the system starts in the basis state $|j\rangle$. Then the probability distribution across states at time t , given a mixed initial state, is simply a weighted average⁷:

$$\begin{aligned} P(t) &= T(t) \cdot P(0) = T(t) \cdot \sum P_j(0) \cdot |j\rangle \\ &= \sum P_j(0) \cdot T(t) \cdot |j\rangle = \sum P_j(0) \cdot P(t; j). \end{aligned}$$

⁷We are slightly abusing our notation here. In this case, $|j\rangle$ refers to the matrix coordinates of the abstract vector: It is a column vector of zeros, except in the j th row, which equals one.

For example, if we set $P_{-1}(0) = P_0(0) = P_{+1}(0) = 1/3$, then the final distribution is just the average of the three distributions produced separately by the basis states $|-1\rangle$, $|0\rangle$, and $|+1\rangle$. This important property of the Markov model will be contrasted with that produced by the quantum model.

Intensity matrix. In many experiments, the evidence is assumed to be sampled from a stationary distribution with a mean rate given by $\lim_{\tau \rightarrow 0} E[X(t + \tau) - X(t) | X(t)] / \tau = \mu$ and dispersion rate given by $\lim_{\tau \rightarrow 0} E[(X(t + \tau) - X(t))^2 | X(t)] / \tau = \sigma^2$. According to the Markov model, $q_{i+1,i} = \lim_{\tau \rightarrow 0} \Pr[X(t + \tau) - X(t) = +\Delta | X(t)] / \tau$, and $q_{i-1,i} = \lim_{\tau \rightarrow 0} \Pr[X(t + \tau) - X(t) = -\Delta | X(t)] / \tau$. This implies that $\mu = (q_{i+1,i} - q_{i-1,i}) \cdot \Delta$ and $\sigma^2 = (q_{i+1,i} + q_{i-1,i}) \cdot \Delta^2$. Solving these two equations for $q_{i+1,i}$ and $q_{i-1,i}$ produces the following solution for the intensities:

$$\begin{aligned} q_{i-1,i} &= \frac{1}{2} \left(\frac{\sigma^2}{\Delta^2} - \frac{\mu}{\Delta} \right), & q_{i+1,i} &= \frac{1}{2} \left(\frac{\sigma^2}{\Delta^2} + \frac{\mu}{\Delta} \right), \\ q_{ii} &= -\frac{\sigma^2}{\Delta^2}, & |i| &< m, \end{aligned}$$

This is called the constant mean drift rate assumption of the Markov model. The rates for the boundary states, $|m\rangle$ and $|-m\rangle$, depend on the type of stopping procedure, which is described next.

Choice measurement. There are two different stopping rules that have been used to measure choice in this task. For the externally controlled procedure, the decision maker has a predetermined amount of time, t^* , to decide. In this case, the level of evidence at the predetermined time is compared to a threshold: if $X(t^*) < \theta$, then a decision in favor of option *A* (noise) is reported; whereas if $X(t^*) > \theta$, then a decision in favor of option *B* (signal) is reported. For the internally controlled procedure, the decision maker decides when to stop. In this case, the process of sampling and accumulating evidence continues until the level of evidence exceeds a threshold magnitude θ . As soon as the threshold is reached, the decision maker stops and makes a choice: option *B* (signal) is chosen if $X(t) = +\theta$; and option *A* (noise) is chosen if $X(t) = -\theta$. For both procedures, the confidence rating corresponding to the decision is assumed to be determined by the magnitude of evidence that exists at the time of the decision.

Externally controlled stopping procedure. For this procedure, a pair of reflecting boundaries is assumed,⁸ so that $q_{-(m-1),-m} = \sigma^2 / \Delta^2 = -q_{-m,-m}$ and $q_{m-1,m} = \sigma^2 / \Delta^2 = -q_{mm}$. In this case, the probability distribution over time, $P(t)$, is computed from Eq. (7b). The reflecting bound model is a positive recurrent system with an equilibrium distribution (Bhattacharya & Waymire, 1990, p. 308). To illustrate the behavior of the model, consider an example

⁸Usually an unrestricted state space, ranging from $-\infty$ to $+\infty$ is assumed for an externally controlled stopping task. But here we are using a state space that is restricted to a finite set. Thus we need to use reflecting bounds to allow the process to continue its evolution after it reaches the finite limits.

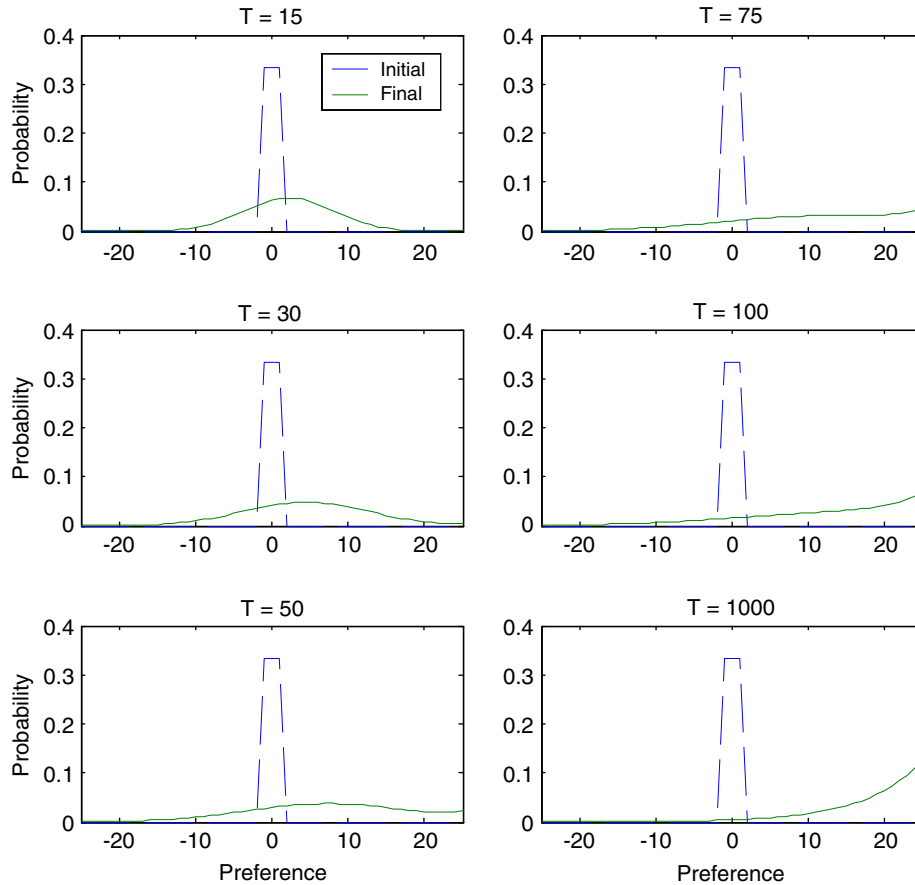


Fig. 3. Distribution over preference states as a function of deliberation time for the Markov random walk model. Horizontal axis indicates confidence level, vertical axis indicates probability, and each panel indicates a different processing time point.

generated by the parameters $\mu = .15$, $\sigma = 1.5$ (for simplicity, we set $\Delta = 1$). Fig. 3 shows snapshots of the distribution across time, with the processing time index shown above each panel. There are two lines within each panel—the dashed line represents the initial distribution, and the solid line represents the distribution produced after processing for a period of time indicated at the top of the panel. As can be seen in this figure, the distribution starts with a peak near zero, and begins moving to the right. The distribution initially becomes normal until it begins to hit the upper (reflecting) boundary. Gradually over time, the distribution builds up near the upper boundary until it finally converges to the equilibrium distribution.

Internally controlled stopping procedure. For this procedure, absorption boundaries are assumed, so that $q_{-mj} = 0 = q_{mj}$. The distribution of choices and response times for the Markov random walk model are derived in Appendix B based on Eq. (7b) (also see Pike, 1966; Shiffrin & Maynard, 1988, for derivations of the moments). To illustrate the behavior of the model, consider an example generated by the parameters $\mu = .3375$, $\sigma = 3$, $\theta = 25$, and the initial distribution was the same as shown in Fig. 3. The joint density and joint cumulative probability distributions for stopping at time t and choosing a response are shown in Fig. 4. The final choice probabilities are .875 and .125 for

correct and incorrect options, and the corresponding mean choice times are 57.83 and 57.76 ms, respectively. As can be seen in this figure, each density has a single mode, and each cumulative distribution has a single inflection point appearing at the beginning of the distribution. This is a general property of the distributions produced by the Markov random walk model.

Confidence ratings. When using the externally controlled stopping procedure, the confidence rating is determined by the probability distribution across evidence levels, conditioned on the decision made at time t^* . More specifically, if a decision-favoring signal is made, then the probability of rating the confidence at level j equals $P_j(t^*) / \sum_{i>\theta} P_i(t^*)$.

When using the internally controlled stopping procedure, a problem arises for the Markov random walk model. In this case, according to this model, the decision maker always stops at the same magnitude of confidence (corresponding to the threshold magnitude θ) on each trial. Thus it is necessary to invent new principles for explaining the distribution of confidence ratings that are empirically observed (see Vickers, 1979, Chapter x, for more discussion of this problem). For example, one could assume that the threshold bound, θ , randomly varies across trials, which would result in variability in the confidence ratings. The quantum model, described next, provides a more elegant way to overcome this problem.

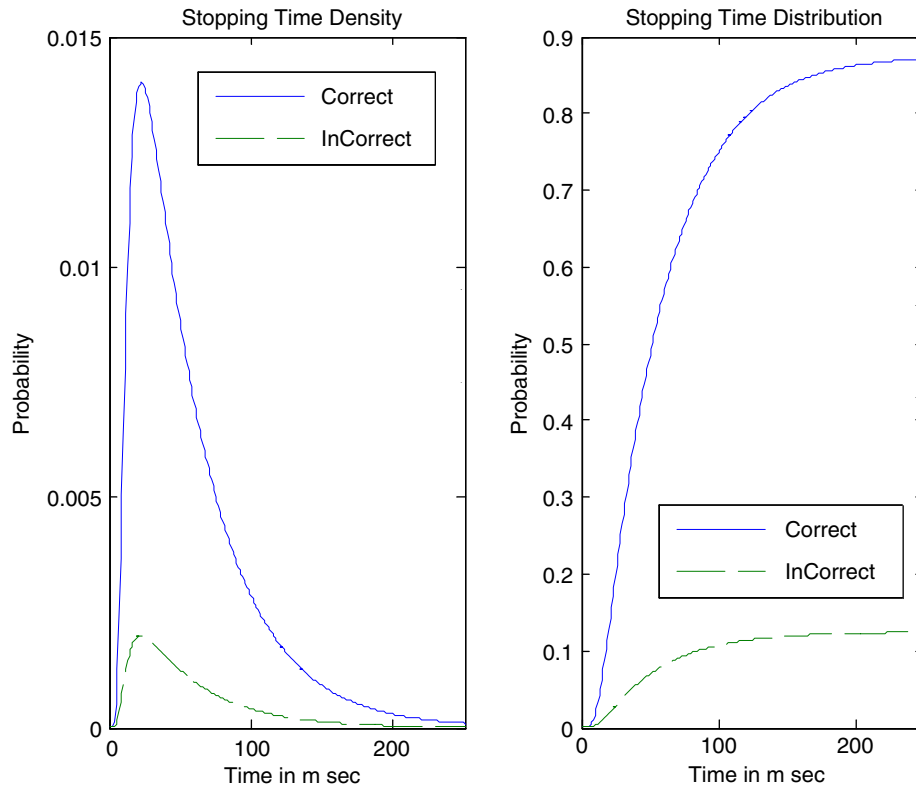


Fig. 4. Joint density and joint cumulative distribution for each response generated by the Markov random walk model.

2.2. Quantum model

Next we present the quantum version of the random walk signal detection model. This is the first attempt to develop such a model, and so we carefully describe each step. First we describe the initial state, followed by a description of the Hamiltonian used to determine the state transitions, and finally we describe the measurement assumptions.

Initial state. The quantum process starts with an initial state $|\psi(0)\rangle = \sum \psi_j(0) \cdot |j\rangle$, where $\psi(0)$ represents the initial distribution of probability amplitudes across the basis states in Ω_{SD} . Similar to the Markov model, when no prior knowledge or information is available, we assume that the probability amplitudes for the initial state are concentrated near the neutral state $|0\rangle$. In particular, for most of the applications presented below, we set $\psi_{-1}(0) = \psi_0(0) = \psi_{+1}(0) = 1/\sqrt{3}$.

The initial distribution for the quantum model has a more complicated effect as compared to the Markov model. Define $\psi(t; j)$ as the probability amplitude distribution across states at time t , given that the system started in the basis state $|j\rangle$. Analogous to the Markov model, we can express the probability amplitude distribution at time t , given an initial superposition state, as a linear combination (see footnote 7):

$$\begin{aligned} \psi(t) &= U(t) \cdot \psi(0) = U(t) \cdot \sum \psi_j(0) \cdot |j\rangle \\ &= \sum \psi_j(0) \cdot U(t) \cdot |j\rangle = \sum \psi_j(0) \cdot \psi(t; j). \end{aligned}$$

However, the probabilities are obtained from $|\psi_i(t)|^2 = |\sum_j \psi_j(0) \cdot \psi_i(t; j)|^2 \neq \sum_j |\psi_j(0)|^2 \cdot |\psi_i(t; j)|^2$, so that interference effects may occur. Later we will illustrate an example of interference effects.

Hamiltonian matrix. The evidence generates a unitary operator that causes the probability amplitude distribution (a quantum wave) to evolve across time to a new state $|\psi(t)\rangle = U(t)|\psi(0)\rangle = \sum \psi_j(t) \cdot |j\rangle$. The vector $\psi(t)$ now represents the distribution of probability amplitudes across the basis states after processing information for a time period equal to t . Analogous to the Markov model, we assume that the evidence is represented by a stationary unitary operator derived from a Hamiltonian of the following simple form:

$$h_{jj} = -(\mu_j/\Delta), \quad h_{j-1j} = h_{j+1j} = -(\sigma_j^2/\Delta^2),$$

where μ_j and σ_j^2 are real numbers. The diagonal parameter, (μ_j/Δ) , is called the *potential* at basis state $|j\rangle$ of the Hamiltonian, and it determines the rate of flow of probability amplitude fed back into a basis state.⁹ The off diagonal parameter, (σ_j^2/Δ^2) , is the diffusion parameter which determines the rate of flow out of a basis state.

⁹Note that adding or subtracting a constant to the diagonal of H has no observable effect: The eigen-decomposition of H can be written $H = VAV$ where V contains the eigenvectors and A the eigenvalues. $H^* = V(A + kI)V = VAV + V(kI)V = H + k \cdot VV = H + (kI)$, where I is the identity matrix. $|e^{-it(kI+H)}\psi|^2 = |e^{-itkI} \cdot e^{-itH}\psi|^2 = |e^{-itk}|^2 |e^{-itH}\psi|^2 = 1 \cdot |e^{-itH}\psi|^2$.

Analogous to the Markov model, we assume that the diffusion rate is constant across columns, so that $\sigma_j^2 = \sigma^2$ for all j . The potential function can be generated from a parametric function defined on the state index j . Recall that in physical mechanics, force is the derivative with respect to position of the potential function, and a linear potential function corresponds to a constant force. Hereafter, we will assume a linear potential function, so that $\mu_j = (\mu \cdot j\Delta)$. This constant force assumption corresponds to the constant mean drift rate assumption of the Markov model. Alternative potential functions, such as for example a quadratic, may prove to be useful and need to be explored in future work.

Choice measurement for externally controlled stopping procedure. According to this procedure, the experimenter determines the amount of time to make the decision, and a choice is expressed at the appointed time, denoted t^* . Therefore a single measurement of the dynamic system is made at time t^* and an observation of either choice A , or choice B , or an indifference response is made at that time. Also a confidence rating for this decision can be observed at this time point. However, we will first discuss the choice measurement, and later we will examine the confidence rating measurement.

According to quantum theory, a measurement is represented by a set of measurement operators, which are applied to the quantum state (see Gardiner, 1991, Chapter 2; Nielsen & Chuang, 2000, Chapter 2). The measurement operator for the choice of option B is the sum of outer products of basis states that exceed the threshold:

$$\mathbf{B} = \sum_{j>\theta} |j\rangle\langle j|.$$

This operator projects the quantum state onto the subspace spanned by the basis vectors $|j\rangle$ for $j>\theta$, representing preferences favoring option B :

$$\mathbf{B}|\psi\rangle = \sum_{j>\theta} |j\rangle\langle j|\psi\rangle = \sum_{j>\theta} \psi_j \cdot |j\rangle.$$

The matrix representation of this projector, with respect to the Ω basis, is the matrix B with $B_{jj} = 1$ for $j>\theta$ and the rest of the elements are zero. The projection resulting from this operator gives the probability amplitudes for choosing B :

$$B \cdot \psi = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \psi_B \end{bmatrix},$$

where ψ_B contains the coordinates of ψ with $j>\theta$.

The measurement operator for the choice of option A is the sum of outer products for the basis states that are less than the threshold:

$$\mathbf{A} = \sum_{j<-\theta} |j\rangle\langle j|.$$

This operator projects the quantum state onto the subspace spanned by the basis vectors $|j\rangle$ for $j<-\theta$ representing preferences favoring option A :

$$\mathbf{A}|\psi\rangle = \sum_{j<-\theta} |j\rangle\langle j|\psi\rangle = \sum_{j<-\theta} \psi_j \cdot |j\rangle.$$

The matrix representation of this projector, with respect to the Ω basis, is the matrix A with $A_{jj} = 1$ for $j<-\theta$ and the rest of the elements are zero. The projection resulting from this operator gives the probability amplitudes for choosing option A :

$$A \cdot \psi = \begin{bmatrix} \psi_A \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix},$$

where ψ_A contains the coordinates of ψ with $j<-\theta$.

The measurement operator for an indifference response N is determined by the operator

$$\mathbf{N} = \mathbf{I} - \mathbf{A} - \mathbf{B} = \sum_{|j|\leq\theta} |j\rangle\langle j|.$$

This operator projects the quantum state onto the subspace spanned by the basis vectors $|j\rangle$ for $|j|\leq\theta$ representing indifference:

$$\mathbf{N}|\psi\rangle = \sum_{|j|\leq\theta} |j\rangle\langle j|\psi\rangle = \sum_{|j|\leq\theta} \psi_j \cdot |j\rangle.$$

The matrix representation of this projector, with respect to the Ω basis, is the matrix N with $N_{jj} = 1$ for $|j|\leq\theta$ and the rest of the elements are zero. The projection resulting from this operator gives the probability amplitudes for choosing indifference:

$$N \cdot \psi = \begin{bmatrix} \mathbf{0} \\ \psi_N \\ \mathbf{0} \end{bmatrix},$$

where ψ_N contains the coordinates of ψ with $|j|\leq\theta$.

Note the sum of all of the measurement operators produces an identity operator: $\mathbf{I} = \mathbf{A} + \mathbf{B} + \mathbf{N}$, or alternatively $A + B + N = I$, where I is the identity matrix. This is called the completeness requirement for the measurement operators.

The probability of observing a choice of option A is obtained by squaring the magnitude of the probability amplitude associated with the event ‘‘Choose A ’’:

$$\begin{aligned} \text{Pr}[\text{Choose } A] &= (A \cdot \psi)^\dagger (A \cdot \psi) \\ &= \psi^\dagger \cdot (A^\dagger \cdot A) \cdot \psi = |\psi_A|^2. \end{aligned} \tag{11a}$$

The probability of observing a choice of option B is defined in a similar manner as

$$\begin{aligned} \text{Pr}[\text{Choose } B] &= (B \cdot \psi)^\dagger (B \cdot \psi) \\ &= \psi^\dagger \cdot (B^\dagger \cdot B) \cdot \psi = |\psi_B|^2. \end{aligned} \tag{11b}$$

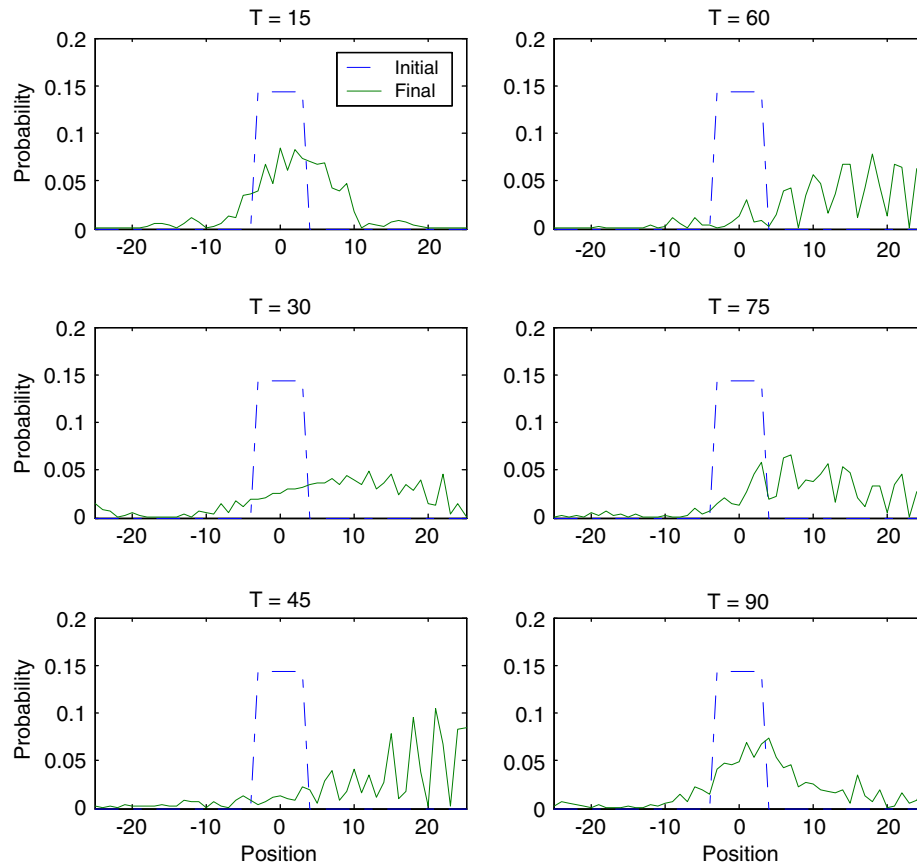


Fig. 5. Distribution over preference states as a function of deliberation time for the quantum random walk model. Horizontal axis indicates confidence level, vertical axis indicates probability, each panel indicates a different processing time point.

The probability of observing an indifference response is given by

$$\begin{aligned} \Pr[\text{Choose } N] &= (N \cdot \psi)^\dagger (N \cdot \psi) \\ &= \psi^\dagger \cdot (N^\dagger \cdot N) \cdot \psi = |\psi_N|^2. \end{aligned} \quad (11c)$$

It follows from the completeness property that $\Pr[\text{Choose } A] + \Pr[\text{Choose } B] + \Pr[\text{Choose } N] = 1.0$.

To grasp how this model works, it is helpful to consider an example. The parameters of the model were set equal to $\mu = 1$, and $\sigma^2 = .50$ ($\Delta = 1$ for simplicity) and Eq. (8b) was used to generate the evolution of the quantum state over time. The results are expressed in terms of the probabilities (squared amplitudes) of the states. Fig. 5 shows snapshots of the probability distributions across time, with the processing time index shown above each panel. There are two lines within each panel—the dashed line represents the initial distribution, and the solid line represents the distribution produced after processing for a period of time indicated at the top of the panel. Comparing Fig. 5 with Fig. 3, it is apparent that the quantum process behaves very differently than the Markov process. In particular, the quantum system produces a wave that tends to slosh back and forth in between the neutral state and the upper boundary. The pattern shown in the figure repeats indefinitely itself as time continuous, and the process never

settles down into a fixed equilibrium distribution. In other words, the quantum process continues to beat against the shore of the upper bound forever. The timing of this oscillation is controlled by the diffusion parameter σ^2 .

Interference effects. Interference is the most fascinating property of quantum dynamics. To examine these effects with the quantum random walk model, consider the effects of using different types of starting superposition states. Given two conflicting hypotheses, there are two very different ways to represent an unbiased initial state. One is to be *indifferent* and place all of the initial distribution near the zero preference state as shown in Fig. 5. But another way is to be *ambivalent*, and start with half of the initial probability distribution placed at the lower extreme (biased in favor of noise), and the other half place at the upper extreme (biased in favor of signal). For example, the operator in a signal detection task may be given prior information by an informant, but the informant may be lying or telling the truth.

This ambivalent initial distribution is illustrated in Fig. 6. The dashed line in the top panel (see the rectangle at the lower bound) represents a starting distribution that is strongly biased in favor of noise. The dashed line in the middle panel (see the rectangle at the upper bound) represents a starting distribution that is strongly biased in favor of signal. The dashed line in the bottom panel

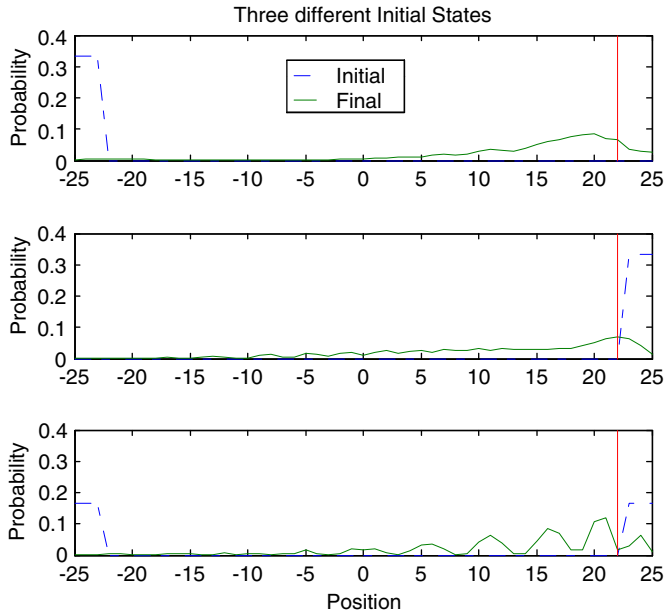


Fig. 6. Interference effect produced by different starting distributions. horizontal axis indicates the confidence level, vertical axis indicates the probability. First panel represents a starting position with confidence favoring option A, middle represents a starting position with confidence favoring option B, the bottom panel represents the ambivalent case where the person starts with some amplitude favoring option A and some probability amplitude favoring option B.

(a rectangle near each bound) has half of the starting distribution place near the lower bound and the other half placed near the upper bound. In this example the parameters were once again set equal to $\mu = 1$, $\sigma^2 = .50$ as before. The solid line within each panel represents the probability distribution that results after processing for a duration of $t = 50$ time units (approximately the first beat against the bound of the oscillating wave).

According to a Markov model, on any given trial, a person starts with either a bias for noise, or else a bias for signal. Thus, the (solid line) probability distribution in the third panel of Fig. 6 should be equal to the average of the two (solid line) distributions produced in the first two panels of Fig. 6. However, as shown in the figure, this is not what happens with the quantum model. In contrast, the quantum model predicts something very different—ambivalence produces two waves coming from opposite directions that crash into each other, causing an interference distribution, which is impossible to get from averaging the two separate distributions.

Suppose that the threshold for responding signal (deciding that an enemy is present, and planning an appropriate attack) is biased toward a high level of confidence ($\theta = 21.5$ on a 25 point scale, see the vertical line in Fig. 6). Then the probability of responding signal (choosing B) in the bottom panel is lower than either the top or middle panels: $\Pr[\text{Choose } B | \text{Start Low or Start High}] = .13 < \min\{\Pr[\text{Choose } B | \text{Start Low}] = .16, \Pr[\text{Choose } B | \text{Start High}] = .18\}$. In other words, it is

impossible to generate the bottom panel from a mixture of the distributions shown in the two top panels.

This interference effect occurs because the probability for choosing B for the bottom panel is determined by (see footnote 7)

$$\begin{aligned} |\psi_B(t)|^2 &= |(1/\sqrt{2})B \cdot U(t) \cdot |-m\rangle + (1/\sqrt{2})B \cdot U(t) \cdot |+m\rangle|^2 \\ &= (1/2)|\psi_B(t; -m) + \psi_B(t; +m)|^2 \\ &\neq (1/2)(|\psi_B(t; -m)|^2 + |\psi_B(t; +m)|^2). \end{aligned}$$

The last line proves that the choice probability produced by the superposition state does not equal the average of the probabilities produced by the two extreme basis states.

Measurement for the internally controlled procedure. According to the quantum model, at any moment in time, the decision maker experiences a superposition of basis states, and thus we cannot say that the decision maker is experiencing any particular confidence level at any moment in time. Thus we cannot use the same stopping rule as used by the Markov model. Instead, we assume that the decision maker makes a series of conscious observations over time, checking periodically whether or not a choice has been made.¹⁰ A measurement is taken by the decision maker at each small time step $\tau > 0$, resulting in an observation of either a choice of A or a choice of B or no choice (cf., Gardiner, 1991, Section 2.3.1). Accordingly, the measurement operators, **A**, **B**, and **N**, are applied after each time step τ , and the probabilities of each response after each time step are obtained from Eqs. (11a–c).

After an observation is made at each time step, the state of the quantum system is projected onto the basis states representing the observed event: If a choice of option A is observed, then the new state is $A \cdot \psi / |A \cdot \psi|$; likewise, if a choice of option B is observed, then the new state is $B \cdot \psi / |B \cdot \psi|$; finally, if a no choice observation is made, the state of the quantum system projects to $N \cdot \psi / |N \cdot \psi|$. Also, if no choice is observed, then the unitary operator (Eq. (8b)) is applied to the resulting no choice state, $N \cdot \psi / |N \cdot \psi|$, for the time period τ , which generates a new state immediately before the next observation is taken.

Choice and stopping time distributions. The distribution of choices and response times for the quantum random walk model are derived in Appendix C. To illustrate the behavior of the model, consider an example generated by the parameters $\mu = 2$, $\sigma^2 = 1$, and $\theta = 20.5$ (on 30 point confidence scale). The same initial distribution as used in Fig. 5. The joint density and joint cumulative probability distributions for stopping at time t and choosing each response are shown in Fig. 7. The final choice probabilities are .86 and .14, and the corresponding mean choice times are 38 and 28 ms, respectively, for correct and incorrect options. As can be seen in this figure, each density is multi-model and each distribution has a multiple inflection

¹⁰One could argue that this corresponds to the cycle time used to simultaneously fire production rules in working memory for cognitive architectures such as Anderson’s ACT-R (Anderson, 1993).

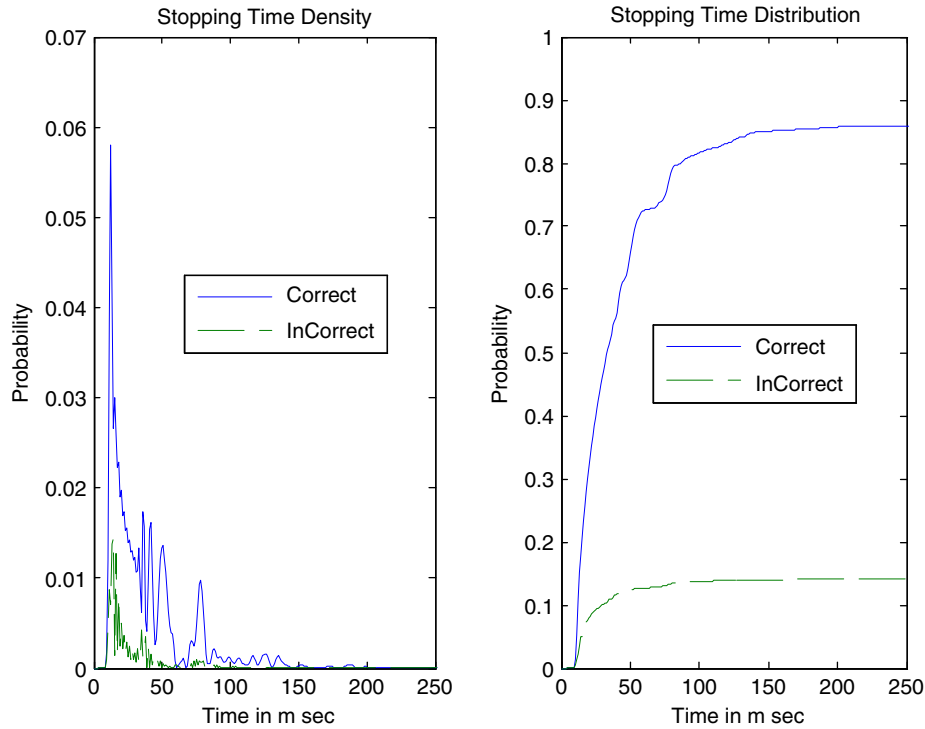


Fig. 7. Joint density and joint cumulative distribution for each response generated by the quantum random walk decision model.

points.¹¹ This is a general property of the quantum random walk model; it is a direct result of the oscillating and beating characteristic of the distributions over time, which was seen in Fig. 5. Recall that the density for the Markov model is generally unimodal. Thus it is clear that the quantum model behaves very differently than Markov model.

Fig. 8 shows how the choice probabilities and mean response times vary as a function of the parameter μ which controls the slope of the potential function. As can be seen in this figure, choice probability increases with μ . Also it is important to note errors are faster than correct responses for relatively high error rates (e.g., conditions emphasizing speed), but errors are slower than correct responses for relatively low error rates (conditions emphasizing accuracy). This is consistent with experimental findings (see Ratcliff & Smith, 2004).

Confidence ratings. One important advantage of the quantum model for signal detection is that it provides a common principle for explaining variation in confidence ratings for both externally and internally controlled stopping tasks. According to the quantum model, if a measurement is taken at t , and a choice is observed, then the quantum state collapses onto the set of confidence levels that exceed the threshold for that choice. A subsequent confidence rating would be based on the probability amplitudes corresponding to the confidence

levels that exceed the threshold (normalized so that the squared amplitudes exceeding the threshold sum to unity). This results in confidence ratings distributed across the range of levels that exceed the threshold. For example, if the threshold is set at $\theta = 20.5$ on a 30 point scale, and a choice favoring option B is observed, then the confidence ratings will be distributed across levels 21 through 30.

More formally, the measurement of confidence is represented by a set of measurement operators, C_j , $j = 1, 2, \dots, m$, one operator for each confidence level. Each of these operators is defined by the two basis states that correspond to same level (disregarding sign) of confidence: $C_j = |-j\rangle\langle -j| + |+j\rangle\langle +j|$. The matrix coordinates of each operator, with respect to the basis set Ω , is a matrix C with zeros everywhere except for the two diagonal elements $C_{+j+j} = C_{-j-j} = 1$.

The joint probability of first making a decision, say signal, and then (immediately) rating the confidence at level j in that decision, is determined by the probability that a choice B is made first, times the probability of rating the confidence at j given that a choice of B has been observed, which equals

$$|B \cdot \psi|^2 \cdot |C_j \cdot \frac{B \cdot \psi}{|B \cdot \psi|}|^2 = |C_j \cdot B \cdot \psi|^2 = |\psi_{+j}|^2.$$

Note that the two matrices commute, $C_j \cdot B = B \cdot C_j$, so that we must predict the same probability when a confidence rating is made first, followed immediately by a decision. Although the latter order of measurement is rarely if ever studied, the current representation of basis states implies that the measurement operators for choice

¹¹The discontinuities in the density are from exact calculations as described in the appendix. They are not noisy estimates from a Monte Carlo simulation.

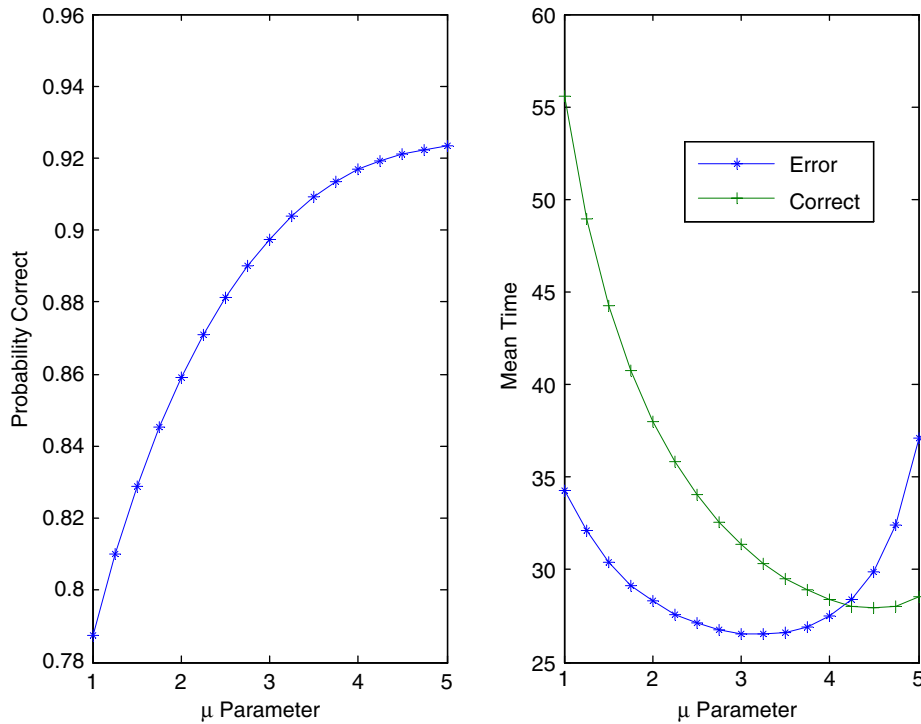


Fig. 8. Choice probability and mean choice time as a function of the parameter μ for the quantum random walk model.

commute with those for confidence rating. Thus the measures are said to be compatible, and order of measurement should not matter. If in fact, order does matter, then this would suggest the need for a different representation in which the measurement operators for choice and confidence do not commute. This would also entail giving up the basis set Ω as the representation, because this basis set uses a common basis state to simultaneously represent a pair of measurement outcomes.

This argument depends on the assumption that a second measurement is taken immediately after the first. If there is some time delay between measurements, then a unitary operator is applied during the delay. In general, the unitary operator does not commute with the measurement operators, so that $|C_j \cdot U \cdot B \cdot \psi|^2 \neq |B \cdot U \cdot C_j \cdot \psi|^2$. Thus order effects can be reintroduced by delays between measurements.

3. Empirical evaluation

3.1. Fits to response time distributions

Response time distributions observed in empirical studies actually do display discontinuities like that shown in Fig. 7. However, researchers usually believe that the discontinuities are caused by sampling error, and so they smooth the empirical distributions to average out the noise. Grouping or smoothing would hide any true discontinuities such as those predicted by the quantum model.

The standard method for evaluating the fit of random walk/diffusion models to response time distributions is to compare the predicted and observed quantiles. Ratcliff and Smith (2004) compared a variety of different random walk/diffusion models in terms of their ability to predict 10 quantiles—that is the 10, 30, 50, 70, and 90th percentiles for the correct and incorrect response time distributions. This method was used below to compare the fits of the quantum versus Markov random walk models to empirical data.

Both models were fit to the two choice probabilities and ten quantiles obtained from two conditions of a signal detection study reported in Ratcliff and Smith (2004). In one condition, speed of responding was emphasized; in another condition, accuracy of responding was emphasized. The three parameters, μ , σ^2 , and T_0 (a motor time component) of the Markov model were estimated separately from each condition; similarly, three parameters, μ , σ^2 , and T_0 of the quantum model were estimated separately for each condition. The parameters were estimated by searching for values that minimized a chi square lack of fit criterion

$$\chi^2 = \sum_k \sum_j \frac{(o_{jk} - p_{jk})^2}{p_{jk}}$$

where o_{jk} and p_{jk} are the observed and predicted proportions, respectively, j indexes the category defined each quantile, and k indexes the response (error vs. correct). The resulting fits are shown in Tables 1 and 2.

Table 1
Data from speed condition of Ratcliff and Smith (2004)

Measure	Error			Correct		
	Obs	Quant	Markov	Obs	Quant	Markov
Probability	0.0830	0.0840	0.0839	0.9170	0.9159	0.9161
Q1	0.3362	0.3292	0.3223	0.3195	0.3192	0.3224
Q2	0.3700	0.3392	0.3526	0.3590	0.3392	0.3528
Q3	0.4171	0.3692	0.3879	0.3951	0.3792	0.3881
Q4	0.4567	0.4092	0.4399	0.4397	0.4392	0.4402
Q5	0.5975	0.5692	0.5515	0.5190	0.5792	0.5519

Table 2
Data from accuracy condition of Ratcliff and Smith (2004)

Measure	Error			Correct		
	Obs	Quant	Markov	Obs	Quant	Markov
Probability	0.0510	0.0952	0.0386	0.9490	0.9047	0.9614
Q1	0.4087	0.3794	0.3754	0.3871	0.3794	0.3758
Q2	0.4647	0.4094	0.4376	0.4473	0.4194	0.4382
Q3	0.5393	0.4494	0.5073	0.5027	0.4794	0.5081
Q4	0.6440	0.5094	0.6084	0.5841	0.5794	0.6093
Q5	0.8432	0.8294	0.8236	0.7680	0.8594	0.8246

Notes: For the Markov model, the following parameters were used: for the speed condition ($\mu = 1.572$, $\sigma = 5.8460$, $T_0 = 28$) producing $\chi^2 = .05$; and for the accuracy condition ($\mu = .8457$, $\sigma = 3.6977$, $T_0 = 28$) producing $\chi^2 = .04$. For the Quantum model the following parameters were used: for the speed condition ($\mu = 4.9439$, $\sigma^2 = 1.2496$, $T_0 = 21.93$) producing $\chi^2 = .1074$; and for the accuracy condition ($\mu = 3.4105$, $\sigma^2 = 1.0345$, $T_0 = 25.94$) producing $\chi^2 = .1074$.

The row labeled “Probability” shows the probability of correct and incorrect responses; and the rows labeled Q1 through Q5 show the five quantiles (expressed in seconds). As can be seen in this table, the Markov model fits the distributions better than the quantum model. Perhaps this should come as no surprise, given that this is the first attempt ever to fit a quantum model to response time distributions. We still have a lot to learn about the quantum dynamic model.

There are two important limitations of the model comparisons reported above. First of all, it is often assumed that the model parameters (e.g., the mean, μ , and the initial starting point) vary across trials (see Ratcliff & Smith, 2004). By adding parameter variability, the Markov model can produce either faster or slower times for incorrect responses. It is interesting to note that the quantum model can achieve this without assuming parameter variability (see Fig. 8). Another important limitation is that the effects of variability caused by non-decision components (e.g., motor response time) were not included in the model fits. The fits reported above simply assumed that the non-decision components contributed a constant additional time (represented by the parameter T_0). For this initial comparison, we wanted to examine the pure predictions of the quantum model in its simplest form. Adding non-decision response time variability could improve the fit but it also would mask the predictions generated by the quantum model.

If there is no striking qualitative evidence for the quantum model, then quantitative fits are not going to convince anyone to move away from the already successful Markov models. What kind of qualitative evidence is needed to encourage continued development of the quantum dynamic model? One important prediction identified in Fig. 7 is that the quantum model generates multi modal response time distributions. In the past, researchers have attempted to explain multi modal distributions in terms of multiple stage processes (Meyer, Yantis, Osman, & Smith, 1985) or multiple decision strategies (Ratcliff, 1988). However, multiple modes may be a signature of a quantum process.

3.2. The disjunction effect

The interference property discussed in conjunction with Fig. 1 is a key prediction of quantum dynamic models. Tversky and Shafir’s (1992; see also Shafir & Tversky, 1992) experimental paradigm on the disjunction effect provides a potentially useful way to investigate this property. These experiments were designed to test an axiom of rational decision making called the sure thing principle (Savage, 1954). According to the sure thing principle, if one prefers option A over B under the state X , and this person also prefers the same option under the complementary state Y , then this person should prefer this same option when it is unknown whether state X or Y

obtains. This principle was tested in an experiment where 98 subjects were told to imagine that they had just played a gamble that gave an even chance of winning \$200 or losing \$100. During the first stage of the task, they were asked to imagine that they had just won the first play, and then they were given a choice whether or not to play a second time. During a second stage a week later they were presented the same problem but asked to imagine that they had lost the first play, and then they were given a choice whether or not to play a second time. During a third stage another week later they were asked to imagine that they did not know the outcome of the first play, and then they were given a choice whether or not to play a second time. The results were that when imagining a win, 69% chose to play again; when imagining a loss, 59% chose to play again; but when the outcome was uncertain, only 36% chose to play again. This pattern of results, called the disjunction effect, violates the sure thing principle of rationality (Savage, 1954).

The disjunction effect was later extended by Shafir and Tversky (1992) using a prisoner's dilemma game (and the findings of the latter study were replicated by Croson, 1999). In this game, each player has two strategies: defect or cooperate. If both players defect, then they lose the most; if both player cooperate, then they both win a modest amount; but if one defects and the other cooperates, then the defector wins the most at the cost of a loss for the cooperator. An experiment was conducted using 80 participants, and each player made a play against three other unknown opponents. In one condition, they were informed that the other player was going to defect; under another condition they were informed that the other player was going to cooperate; and under a third condition, the other player's strategy was unknown and uncertain. When informed about the other opponent's strategy, the participants rarely ever cooperated (less than 16%). However, when the other player's strategy was unknown, the rate of cooperation increased to 37%, which is higher than either of the two previous conditions. In this case, it was advantageous not to know the other player's strategy because everyone gained by cooperation.¹²

Quantum dynamics may provide a way to account of these paradoxical results. Consider Fig. 1 again, but now interpret the initial state, X_0 , as the players initial disposition to defect before thinking about the opponent; interpret the two hidden states as the two possible dispositions that result after the player thinks about the opponent's moves ($X_1 \rightarrow$ opponent cooperates, $X_2 \rightarrow$ opponent defects); the final state X_4 is used to determine the player's observed choice to defect. In the first case, the player knows for sure the opponents move; on half the trials, the player is told that the opponent will defect, and on the other half, the player is told that the opponent will cooperate. This corresponds to the case in which the

photon's path is known, so the probability of defecting is the sum of the probabilities for each path separately:

$$(|\langle X_4|X_1\rangle\langle X_1|X_0\rangle|^2 + |\langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2).$$

In the second case, the player does not know what the opponent will do: it is uncertain whether the opponent could defect or cooperate. This corresponds to the case in which the photon's path is unknown: so the probability of defecting is affected by interference:

$$|\langle X_4|X_1\rangle\langle X_1|X_0\rangle + \langle X_4|X_2\rangle\langle X_2|X_0\rangle|^2.$$

Recall that the interference effect depends on the angle between the probabilities amplitudes that are being summed together. According to this quantum interpretation, the angle is a function of the information about the opponent's move. Variation on the amount of information given to the player is therefore predicted to change the angle, and thus systematically change the size of the disjunction effect. In this way, the quantum model makes specific quantitative predictions for the size of the disjunction effect.

4. Extensions and conclusions

Summary. This article accomplished several goals. First, we developed a general quantum random walk model, and directly compared its assumptions with a general Markov random walk model. Second, we developed a specific application of the quantum model to signal detection, and derived the choice probability and response time distributions for this model. Third, we empirically compared the quantum model with a Markov model in terms of its ability to fit choice and response time distributions observed in a signal detection task. Fourth, we indicated how the quantum model could help explain some paradoxical findings from the field of decision making such as the disjunction effect. Now we wish to point out some other possible applications and extensions of the quantum model, and review some related work on this new topic.

Extensions. We explored just one possible type of application of quantum dynamics to psychology, and many more can easily be imagined. Consider, for example, the Markov random walk model of preference evolution developed by Falmagne et al. (1997) and Regenwetter et al. (1999). This theory was used to describe how an individual's preferences for political candidates evolve over time in response to political information from the environment. According to the Markov version of this theory, the state of preference of an individual at any moment in time is represented by a partial order over candidates, and information changes this state by moving a person from one partial order to another. This theory could be reformulated as a quantum random walk process. However, this would require one to give up traditional views about preference states. According to a quantum version of this theory, each possible partial order represents a basis, and an individual's current state is a

¹²Shafir and Tversky (1992) point out that Newcomb's paradox can also be interpreted as an example of a disjunction effect. They also hint at a quantum interpretation here.

superposition of these basis states. Information transforms this superposition state by applying a unitary operator to the state. Only when a preference is revealed does the state collapse onto a particular partial order. Thus preferences do not jump from one partial order to another over time, and instead all partial orders coexist during information processing. No particular partial order exists until the preferences are revealed. This is a new and radical way to view preference states.

Related research. Several other applications of quantum principles to psychological research have appeared recently. Bordley (1998; Bordley & Kadane, 1999) has used quantum probability rules to explain paradoxes in human probability judgments. Lambert-Mogiliansky, Zamir, and Zwirn (2004) have used non-commutative measurement operators to explain order effects on human judgments. Gabora and Aerts (2002) used entangled quantum states to explain the contextual manner in which concepts are evoked. Eisert, Wilkens, and Lewenstein (1999; see also Piotrowski, & Sladkowski, 2003) have proposed quantum strategies to explain cooperation in prisoner's dilemma games. Grossberg (2000) argues that the organization of the brain obeys the same uncertainty and complementary principles as does the physical world with which it interacts. Pribram (1993) provides a compilation work on quantum neural networks. Finally, an ambitious attempt to explain consciousness in terms of quantum and classical interactions was proposed by Penrose (1989). Some other scholars of consciousness also have employed a quantum approach (e.g., Hameroff, 1994; Miranker, 2002; Stapp, 1999; Woolf & Hameroff, 2001).

Concluding comments. We are only just beginning to understand the theoretical potential of quantum dynamic models, and new applications need to be explored. Markov models of information processing (Townsend & Ashby, 1983) could be reformulated according to quantum principles. Quantum computing theory (Nielsen & Chuang, 2000) may provide a whole new way to synthesize production rule systems (Anderson, 1993) and parallel processing systems (Rumelhart & McClelland, 1986). However, zealous advocates of quantum models must cautiously seek more empirical evidence to justify the additional complexity of this theory.

Acknowledgments

Part of this work was presented at the North American Falmagne Festschrift, University of Michigan, 2004. This work was supported by NSF Methodology, Measurement, Statistics SES-0083511 Biocomplexity and by NIMH R01 MH68346-02. We thank Ariane Lambert-Mogiliansky and Philip Smith for comments.

Appendix A. Linear algebra in Hilbert spaces

Quantum theory works with abstract vectors and operations on vectors defined within a Hilbert space. This

appendix is limited to finite dimensional vector spaces. A finite dimensional Hilbert space is an abstract vector space defined on a complex field which is equipped with an inner product. (Generally, it is also required that the space is complete, but this is trivial in this finite dimension case.) Let $|\psi\rangle$ and $|\chi\rangle$ be two arbitrary vectors in the Hilbert space of dimension m . The inner product between these two vectors is symbolized as $\langle\chi|\psi\rangle$.

The inner product satisfies three properties: positive semi-definiteness $\langle\psi|\psi\rangle \geq 0$, anti-symmetry $\langle\chi|\psi\rangle = \langle\psi|\chi\rangle^*$, and linearity $\langle\psi|(a\chi_1 + b\chi_2)\rangle = a\langle\psi|\chi_1\rangle + b\langle\psi|\chi_2\rangle$. In general, the inner product is a complex number so that $\langle\chi|\psi\rangle = (\alpha + i\beta) = \rho \cdot e^{i\phi} = \rho \cdot [\cos(\phi) + i\sin(\phi)]$, where $i = \sqrt{-1}$, $\rho^2 = |\langle\chi|\psi\rangle|^2 = \langle\chi|\psi\rangle^* \langle\chi|\psi\rangle = (\alpha - i\beta) \cdot (\alpha + i\beta) = \alpha^2 + \beta^2$, $\phi = \tan^{-1}(\beta/\alpha)$. The complex conjugate is defined as $\langle\chi|\psi\rangle^* = (\alpha - i\beta)$. The square of the sum of two inner products can be expanded as follows: $|\langle\phi|\chi\rangle + \langle\chi|\psi\rangle|^2 = [\langle\phi|\chi\rangle + \langle\chi|\psi\rangle] \cdot [\langle\phi|\chi\rangle + \langle\chi|\psi\rangle]^* = [\langle\phi|\chi\rangle + \langle\chi|\psi\rangle] \cdot [\langle\phi|\chi\rangle^* + \langle\chi|\psi\rangle^*] = \langle\phi|\chi\rangle \langle\phi|\chi\rangle^* + \langle\chi|\psi\rangle \langle\chi|\psi\rangle^* + \langle\chi|\psi\rangle \langle\phi|\chi\rangle^* + \langle\phi|\chi\rangle \langle\chi|\psi\rangle^*$. Defining $\langle\phi|\chi\rangle = (\alpha_1 + i\beta_1)$ and $\langle\chi|\psi\rangle = (\alpha_2 + i\beta_2)$, then $\langle\chi|\psi\rangle \langle\phi|\chi\rangle^* + \langle\phi|\chi\rangle \langle\chi|\psi\rangle^* = (\alpha_2 + i\beta_2) (\alpha_1 - i\beta_1) + (\alpha_1 + i\beta_1) (\alpha_2 - i\beta_2) = (\alpha_1 \cdot \alpha_2 + \beta_1 \cdot \beta_2) = 2 \cdot |\langle\chi|\psi\rangle| \cdot |\langle\phi|\chi\rangle| \cdot \cos(\theta)$, with

$$\cos(\theta) = \frac{\alpha_1 \alpha_2 + \beta_1 \beta_2}{\sqrt{\alpha_1^2 + \beta_1^2} \sqrt{\alpha_2^2 + \beta_2^2}}.$$

Any set of m linearly independent vectors defines a basis for the Hilbert space. Any point in the Hilbert space can be reached by a linear combination of the basis vectors. However, it is convenient to choose basis vectors to be orthonormal: $\langle i|j\rangle = 1$ if $i = j$ and zero otherwise. It is always possible to transform a linearly independent set of basis vectors into an orthonormal set by a procedure called the Gram Schmidt procedure. Define the set $\Omega = \{|i\rangle$ for $i = 1, 2, \dots, m\}$ as the orthonormal basis for the Hilbert space. The projection of a vector $|\psi\rangle$ on to the basis vector $|i\rangle$ is given by $\langle i|\psi\rangle \cdot |i\rangle$. Thus we can define the coordinates of a vector $|\psi\rangle$ in terms of the projections onto the basis vectors as

$$|\psi\rangle = \sum \langle i|\psi\rangle \cdot |i\rangle = \sum \psi_i \cdot |i\rangle,$$

where ψ_i is the i th coordinate of $|\psi\rangle$ represented in terms of the Ω basis. The column vector ψ has ψ_i as its coordinate in the i th row.

An operator \mathbf{L} transforms one state into another: $|\psi\rangle = \mathbf{L}|\chi\rangle$. A linear operator satisfies the property: $\mathbf{L}(\alpha \cdot |\psi\rangle + \beta \cdot |\chi\rangle) = \alpha \cdot \mathbf{L}|\psi\rangle + \beta \cdot \mathbf{L}|\chi\rangle$. We can express a linear transformation in terms of the Ω basis as

$$\begin{aligned} \psi_i &= \langle i|\psi\rangle = \langle i|\mathbf{L}|\chi\rangle = \langle i|\mathbf{L} \sum \chi_j \cdot |j\rangle \\ &= \langle i|\mathbf{L} \sum |j\rangle \langle j|\chi\rangle. \end{aligned}$$

Assuming the linear property, we can rewrite this as

$$= \sum \langle i|\mathbf{L}|j\rangle \langle j|\chi\rangle = \sum L_{ij} \cdot \chi_j,$$

Finally, with respect to the Ω basis, we have $\psi = L \cdot \chi$, where L is a matrix with element $L_{ij} = \langle i | L | j \rangle$ in row i column j .

The inner product, $\langle \chi | \psi \rangle$, can be viewed as a linear transformation of the vector $|\psi\rangle$. In this view, $\langle \chi |$ acts as a linear operator on $|\psi\rangle$ that maps it into a scalar. The matrix coordinates of this linear operator form a row vector denoted, χ^\dagger , with coordinates $\chi = [\chi_1^* \chi_2^* \dots \chi_m^*]$. Therefore, the inner product of two vectors $|\chi\rangle$ and $|\psi\rangle$ can be computed using the coordinates of the Ω basis as $\langle \chi | \psi \rangle = \chi \cdot \psi = \sum \chi_i^* \cdot \psi_i$.

The outer product $|i\rangle\langle i|$ of the basis vector $|i\rangle$ is a linear operator that projects an arbitrary vector $|\psi\rangle$ onto the subspace of the vector $|i\rangle$ as follows: $|i\rangle\langle i | \psi \rangle = \psi_i \cdot |i\rangle$. It satisfies the idempotent property $|i\rangle\langle i|^2 = |i\rangle\langle i|$, and the set of these operators satisfies the completeness property $|\psi\rangle = \sum \psi_i \cdot |i\rangle = \sum |i\rangle\langle i | \psi \rangle = \mathbf{I} | \psi \rangle \rightarrow \sum |i\rangle\langle i| = \mathbf{I}$, where \mathbf{I} is the identity operator $\mathbf{I} | \psi \rangle = | \psi \rangle$ for any vector.

For each vector $|\psi\rangle$, we have the inner product $\langle \psi | \psi \rangle = \sum \psi_i^* \cdot \psi_i$, and the row vector $\psi = [\psi_1^* \dots \psi_m^*]$, corresponding to this inner product, can be used to define an abstract dual vector $\langle \psi | = \sum \psi_i^* \cdot \langle i |$. For each linear transformation $|\chi\rangle = \mathbf{L} | \psi \rangle$, we can define a corresponding adjoint transformation $\langle \chi | = \langle \psi | \mathbf{L}$. If L is the matrix representation of \mathbf{L} with element L_{ij} in row i column j , then its adjoint is defined as the matrix L^\dagger with element L_{ji} in row j column i equal to $L_{ji} = L_{ij}^*$. The inner product between two vectors can be rewritten using the adjoint notation: $\langle \chi | \psi \rangle = \chi \psi = \sum \chi_i^* \cdot \psi_i$. A Hermitian matrix H has the property that $H = H^\dagger$.

Hermitian matrices permit a spectral decomposition: $H = V A V$ where $V V^\dagger = I$ and A is diagonal. V is a matrix with orthonormal eigenvectors in the columns, and the elements λ_j of A are real numbers. A matrix function of a Hermitian matrix is defined as $f(H) = V f(A) V$ where $f(A)$ is a diagonal matrix with $f(\lambda_j)$ as the j th diagonal element.

A unitary matrix U has the property $U U^\dagger = U^\dagger U = I$, where I is the identity matrix. Unitary operators preserve inner products: If $\psi = U \chi$ then $\psi \psi = \chi U^\dagger U \chi = \chi I \chi = \chi \chi$. This property is important because state vectors in quantum dynamics are required to be unit length: $\psi \psi = 1$.

Any unitary matrix U can be expressed in terms of another matrix, called the Hamiltonian matrix, denoted H , where H is Hermitian. The relation between the two is given by the matrix functions: $H = i \cdot \ln(U)$ and $U = e^{-i \cdot H}$. The Hamiltonian matrix, H , must be Hermitian in order to guarantee that that $e^{-i \cdot H}$ is unitary.

Appendix B. Markov model

For the Markov version of the random walk model, states $|-m\rangle$ and $|m\rangle$ are absorbing states for choosing options A and B , respectively, and the remaining states $|j\rangle$ for $j = -(m-1)$ to $(m-1)$ are transient states. Recall that $P(t)$ is the column vector containing the probability distribution over all the states at time t (which sums to one). The matrix $T(t)$ represents the probability of

transiting to a row state from a column state after a period of time t , and Q represents the rates of these transitions. The probability of making a transition from one state to another, disregarding time, are based on the matrix W , with elements

$$w_{ij} = \frac{q_{ij}}{-q_{ii}} \text{ for } i \neq j \text{ and } w_{ij} = 0 \text{ for } (i,j) \in \Omega^2.$$

The derivation of distribution of choice times requires working with matrices restricted to the transient states. Define $\bar{P}(t)$ as the probability distribution over the transient states at time t (which sums to less than one for $t > 0$). Define $\bar{T}(t)$ as the transition matrix obtained by deleting the first and last, rows and columns, of $T(t)$, so that it contains only the transitions among transient states. Similarly, \bar{Q} and \bar{W} are obtained by deleting the first and last, rows and columns, of Q and W respectively. Define V as a matrix with two rows, where the first row, denoted V_A , is defined by columns $-(m-1)$ to $(m-1)$ in the first row of W , and the second row, denoted V_B , is defined by columns $-(m-1)$ to $(m-1)$ in the last row of W . In other words, V contains the probabilities of transiting from a transient state to an absorbing state.

First we simply state the formula for the choice probabilities (see, e.g., chapters 2 and 3 of [Bhattacharya and Waymire, 1990](#)). Define Π as a matrix with two rows, the first row, Π_A , containing the probabilities of choosing option A (e.g., enemy); and the second row, Π_B , containing the probabilities of choosing option B (e.g., friendly), where each column is the probability conditioned on the initial state, $X(0) = j, j \in \{-(m-1), \dots, (m-1)\}$. These choice probabilities are obtained from the matrix product:

$$\Pi = V \cdot (I - \bar{W})^{-1}. \tag{B.1}$$

Post multiplying by the initial state distribution gives the probability of choosing option A , $\pi_A = \Pi_A \cdot P(0)$, and the probability of choosing option B , $\pi_B = \Pi_B \cdot P(0)$. The mean response times are computed from $(\lambda = (\sigma/A)^2 = -q_{ii})$

$$MT = \lambda^{-1} \cdot V \cdot (I - \bar{W})^{-2} \cdot P(0). \tag{B.2}$$

The mean time to choose A is $MT_A = MT_1/\pi_A$, and the mean time to choose B is $MT_B = MT_2/\pi_B$ where MT_1 and MT_2 are the first and second elements in the row vector MT .

Next we derive the distribution of choice response times. Consider the random walk process prior to absorption, which is restricted to the transient states. It follows from the strong Markov property that the transition probabilities of the restricted process obey the forward Kolmogorov equation ([Bhattacharya and Waymire, 1990, p. 319](#)):

$$\frac{d}{dt} \bar{T}(t) = \bar{Q} \cdot \bar{T}(t) \text{ with the solution } \bar{T}(t) = e^{\bar{Q} \cdot t}. \tag{B.3}$$

Note that the probability of choosing option B is the sum of the probabilities for two mutually exclusive events:

Event E_1 occurs when the process absorbs at state $|m\rangle$ before time t ; Event E_2 occurs when the process remains within the transient states before time t but eventually absorbs at state $|m\rangle$ after time t . The probability of event E_1 is the joint probability of stopping before time t and choosing option B , which is the cumulative distribution that we are trying to find. The probability of event E_1 is denoted $F(t \wedge B)$. The probability of event E_2 equals the joint probability of (1) starting at an initial state, (2) moving to a new transient state during the period of time t , and then (3) finally choosing option B from the new transient state. The probability of event E_2 is equal to $\Pi_B \cdot \bar{T}(t) \cdot \bar{P}(0)$. Therefore $\pi_B = F(t \wedge B) + \Pi_B \cdot \bar{T}(t) \cdot \bar{P}(0)$, which implies

$$F(t \wedge B) = \pi_B - \Pi_B \cdot \bar{T}(t) \cdot \bar{P}(0). \quad (\text{B.4})$$

The joint probability density of choosing option B at time t is obtained by taking the first derivative with respect to time:

$$f(t \wedge B) = (d/dt)F(t \wedge B) = \Pi_B \cdot (-\bar{Q}) \cdot \bar{T}(t) \cdot \bar{P}(0). \quad (\text{B.5})$$

The cumulative joint probability of stopping at time t and choosing A is

$$F(t \wedge A) = \pi_A - \Pi_A \cdot \bar{T}(t) \cdot \bar{P}(0). \quad (\text{B.6})$$

The joint probability density of choosing option B at time t is obtained by

$$f(t \wedge A) = (d/dt)F_A(t) = \Pi_A \cdot (-\bar{Q}) \cdot \bar{T}(t) \cdot \bar{P}(0). \quad (\text{B.7})$$

Eqs. (B.1)–(B.7), are quickly and easily computed using the matrix operators commonly available in matrix programming languages such as Matlab or Mathematica or Gauss.

Appendix C. Quantum model

For the quantum random walk model, we assume that the choice system does not start in states $|j\rangle$ for $|j| > \theta$. In other words, for $|j| > \theta$ we have $\psi_{-j}(0) = \psi_j(0) = 0$ in the initial state, $\psi(0)$. Thus the probability of observing no choice at the beginning of the experiment ($n = 0$) is assumed to be one, i.e., $|\gamma(0)|^2 = 1$.

Starting at $|\psi(0)\rangle$, the probability amplitude for passing through U for a period of time τ , and then observing a choice of option A in the first time step ($n = 1$) equals

$$\alpha(1) = A \cdot U(\tau) \cdot \psi(0).$$

The corresponding probability amplitude for observing a choice of B equals

$$\beta(1) = B \cdot U(\tau) \cdot \psi(0).$$

The probability amplitude of no choice being observed in the first step is given by

$$\gamma(1) = N \cdot U(\tau) \cdot \psi(0).$$

If no choice is observed during the first time step, then the quantum state is projected onto the subspace generated

by the no choice operator:

$$\psi(1) = \frac{N \cdot U(\tau) \cdot \psi(0)}{|N \cdot U(\tau) \cdot \psi(0)|} = \frac{\gamma(1)}{|\gamma(1)|}.$$

Given that no choice was observed in the first step, the probability amplitude for passing through U on the second step ($n = 2$), and then choosing option A equals

$$\alpha(2) = A \cdot U(\tau) \cdot \psi(1).$$

The joint probability of not making a choice on the first step and then observing a choice for A on the second step is then given by

$$\begin{aligned} p_A(2\tau) &= |\alpha(2)|^2 \cdot |\gamma(1)|^2 \\ &= |A \cdot U(\tau) \cdot \psi(1)|^2 \cdot |\gamma(1)|^2 \\ &= \frac{|A \cdot U(\tau) \cdot N \cdot U(\tau) \cdot \psi(0)|^2}{|\gamma(1)|^2} \cdot |\gamma(1)|^2, \end{aligned}$$

which simplifies to

$$p_A(2\tau) = |A \cdot U(\tau) \cdot N \cdot U(\tau) \cdot \psi(0)|^2.$$

Given that no choice was observed in the first step, the probability amplitude for passing through U on the second step ($n = 2$), and then choosing option B equals

$$\beta(2) = B \cdot U(\tau) \cdot \psi(1).$$

The corresponding joint probability of not choosing on the first step and then observing a choice of B is

$$p_B(2\tau) = |B \cdot U(\tau) \cdot N \cdot U(\tau) \cdot \psi(0)|^2.$$

Given no choice on the first step, the probability amplitude for a no choice observation after the second step is given by

$$\gamma(2) = N \cdot U(\tau) \cdot \psi(1).$$

If no choice is observed during the first two steps, then the quantum state is projected onto the subspace generated by the no choice operator:

$$\begin{aligned} \psi(2) &= \frac{N \cdot U(\tau) \cdot \psi(1)}{|\gamma(2)|} = \frac{N \cdot U(\tau) \cdot N \cdot U(\tau) \cdot \psi(0)}{|\gamma(2)| \cdot |\gamma(1)|} \\ &= \frac{[N \cdot U(\tau)]^2 \cdot \psi(0)}{|\gamma(2)| \cdot |\gamma(1)|}. \end{aligned}$$

The preceding steps suggest the following generalization: The probability of not observing a choice in the $n-1$ steps, and passing through U for another step, and then choosing option A equals

$$p_A(n\tau) = |A \cdot U(\tau) \cdot [N \cdot U(\tau)]^{n-1} \cdot \psi(0)|^2.$$

The corresponding probability for observing a choice of B is

$$p_B(n\tau) = |B \cdot U(\tau) \cdot [N \cdot U(\tau)]^{n-1} \cdot \psi(0)|^2.$$

If no choice is observed during the first two steps, then the quantum state is projected onto the subspace generated

by the no choice operator:

$$\psi(n) = \frac{N \cdot U(\tau) \cdot \psi(n-1)}{|\gamma(n)|} = \frac{[N \cdot U(\tau)]^n \cdot \psi(0)}{\prod_{j=1}^n |\gamma(j)|}.$$

This generalization can be proven by induction. Consider the probability of observing a choice for option A on the next step. Given that no choice was observed during the first n time steps, the probability amplitude for passing through \mathbf{U} and then observing a choice for A is given by

$$\alpha(n) = A \cdot U(t) \cdot \psi(n).$$

The joint probability of not observing a choice in the first n steps, and passing through \mathbf{U} for another step ($n+1$), and then choosing option A equals

$$\begin{aligned} p_A((n+1) \cdot \tau) &= |A \cdot U(\tau) \cdot \psi(n)|^2 \cdot \prod_{j=1}^n |\gamma(j)|^2 \\ &= \frac{|A \cdot U(\tau) \cdot [N \cdot U(\tau)]^n \cdot \psi(0)|^2}{\prod_{j=1}^n |\gamma(j)|^2} \cdot \prod_{j=1}^n |\gamma(j)|^2 \\ &= |A \cdot U(\tau) \cdot [N \cdot U(\tau)]^n \cdot \psi(0)|^2. \end{aligned}$$

If no choice is observed during the first $n+1$ steps, then the quantum state is projected onto the subspace generated by the no choice operator:

$$\psi(n+1) = \frac{N \cdot U(\tau) \cdot \psi(n)}{|\gamma(n+1)|} = \frac{[N \cdot U(\tau)]^{n+1} \cdot \psi(0)}{\prod_{j=1}^{n+1} |\gamma(j)|}.$$

Note that this is consistent with the advancement of the earlier solution by one step, and therefore, this completes the inductive proof.

Recall that the unitary matrix can be expressed in terms of the Hamiltonian using the matrix exponential form: $U(\tau) = e^{-iH \cdot \tau}$. Substituting this for the unitary matrix in the previous solutions produces the final solutions for options A and B respectively. The probability of stopping and choosing option A at time $t = n \cdot \tau$ is

$$p_A(n\tau) = |A \cdot e^{-iH \cdot \tau} \cdot [N \cdot e^{-iH \cdot \tau}]^{n-1} \cdot \psi(0)|^2.$$

The probability of stopping and choosing option B at time $t = n \cdot \tau$ is

$$p_B(n\tau) = |B \cdot e^{-iH \cdot \tau} \cdot [N \cdot e^{-iH \cdot \tau}]^{n-1} \cdot \psi(0)|^2.$$

Appendix D. Continuous state models

Markov model. The Kolmogorov equation (Eq. (4b)) for the Markov random walk model can be rewritten in terms of the individual equations as follows:

$$\frac{d}{dt} P_j(t) = q_{j,j-1} P_{j-1}(t) + q_{jj} P_j(t) + q_{j,j+1} P_{j+1}(t).$$

Inserting the following specifications for the intensities from Section 2.1

$$\begin{aligned} q_{j,j-1} &= \frac{1}{2} \left(\frac{\sigma^2}{\Delta^2} + \frac{\mu}{\Delta} \right), & q_{j,j+1} &= \frac{1}{2} \left(\frac{\sigma^2}{\Delta^2} - \frac{\mu}{\Delta} \right), \\ q_{jj} &= -\frac{\sigma^2}{\Delta^2} \end{aligned}$$

produces the following equation:

$$\begin{aligned} \frac{d}{dt} P_j(t) &= \frac{1}{2} \left(\frac{\sigma^2}{\Delta^2} + \frac{\mu}{\Delta} \right) \cdot P_{j-1}(t) \\ &\quad - \frac{\sigma^2}{\Delta^2} \cdot P_j(t) + \frac{1}{2} \left(\frac{\sigma^2}{\Delta^2} - \frac{\mu}{\Delta} \right) \cdot P_{j+1}(t). \end{aligned}$$

If we assume that $x = j\Delta$ and that the adjacent basis states differ by a small increment Δ , then the equation can be rewritten as

$$\begin{aligned} \frac{d}{dt} P(t, x) &= \frac{1}{2} \frac{\sigma^2}{\Delta^2} \{ [P(t, x + \Delta) - P(t, x)] \\ &\quad - [P(t, x) - P(t, x - \Delta)] \} \\ &\quad - \frac{1}{2} \frac{\mu}{\Delta} [P(t, x + \Delta) - P(t, x - \Delta)] \\ &= \frac{\sigma^2}{2} \frac{P(t, x + \Delta) - P(t, x) - P(t, x) + P(t, x - \Delta)}{\Delta} \\ &\quad - \mu \cdot \frac{[P(t, x + \Delta) - P(t, x)] + [P(t, x) - P(t, x - \Delta)]}{2\Delta} \end{aligned}$$

and as $\Delta \rightarrow 0$ we obtain in the limit

$$\frac{\partial}{\partial t} P(t, x) = \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} P(t, x) - \mu \frac{\partial}{\partial x} P(t, x).$$

Quantum model. The Schrödinger equation (Eq. (8b)) for the quantum random walk model can be rewritten in terms of the individual equations as follows:

$$i \frac{d}{dt} \psi_j(t) = h_{j,j-1} \cdot \psi_j(t) + h_{jj} \cdot \psi_j(t) + h_{j,j+1} \cdot \psi_j(t).$$

Inserting the specifications from Section 2.2

$$h_{jj} = \mu_j, \quad h_{j-1,j} = -(\sigma^2/\Delta^2) = h_{j+1,j},$$

and $h_{ij} = 0$ for $|i-j| > 1$,

yields

$$\begin{aligned} i \frac{d}{dt} \psi_j(t) &= \left[-\frac{\sigma^2}{\Delta^2} \cdot \psi_{j-1}(t) + \mu_j \cdot \psi_j(t) \right. \\ &\quad \left. - \frac{\sigma^2}{\Delta^2} \cdot \psi_{j+1}(t) \right] \\ &= \left(\mu_j - 2 \frac{\sigma^2}{\Delta^2} \right) \cdot \psi_j(t) \\ &\quad - \frac{\sigma^2}{\Delta^2} \{ [\psi_{j+1}(t) - \psi_j(t)] \\ &\quad - [\psi_j(t) - \psi_{j-1}(t)] \}. \end{aligned}$$

Without loss in generality (see footnote 9), we can define $\mu_j = v_j + 2(\sigma^2/\Delta^2)$ in which case we the above equation simplifies as follows:

$$i \frac{d}{dt} \psi_j(t) = v_j \cdot \psi_j(t) - \frac{\sigma^2}{\Delta^2} [\psi_{j+1}(t) - \psi_j(t) - [\psi_j(t) - \psi_{j-1}(t)]]$$

If we assume that $x = j\Delta$, where the adjacent basis states differ by a small increment then the differential equation can be rewritten as

$$\begin{aligned} i \frac{d}{dt} \psi(t, x) &= v(x) \cdot \psi(t, x) - \frac{\sigma^2}{\Delta^2} \cdot [\psi(t, x + \Delta) \\ &\quad - \psi(t, x)] - [\psi(t, x) - \psi(t, x - \Delta)] \\ &= v(x) \cdot \psi(t, x) \\ &\quad - \sigma^2 \cdot \frac{(\psi(t, x + \Delta) - \psi(t, x)) - (\psi(t, x) - \psi(t, x - \Delta))}{\Delta} \end{aligned}$$

and as $\Delta \rightarrow 0$, we obtain in the limit

$$i \cdot \frac{\partial}{\partial t} \psi(t, x) = -\sigma^2 \frac{\partial^2}{\partial x^2} \psi(t, x) + v(x) \cdot \psi(t, x).$$

References

- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Batchelder, W. H., & Reiffer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, 6, 57–86.
- Bhattacharya, R. N., & Waymire, E. C. (1990). *Stochastic processes with applications*. NY: Wiley.
- Bockholt, U. (2001). A Thurstonian analysis of preference change. *Journal of Mathematical Psychology*, 46, 300–314.
- Bordley, R. F. (1998). Quantum mechanical and human violations of compound probability principles: Toward a generalized Heisenberg uncertainty principle. *Operations Research*, 46, 923–926.
- Bordley, R. F., & Kadane, J. B. (1999). Experiment dependent priors in psychology and physics. *Theory and Decision*, 47, 213–227.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic cognition approach to decision making. *Psychological Review*, 100, 432–459.
- Cox, D. R., & Miller, H. D. (1965). *The theory of stochastic processes*. London: Chapman Hall.
- Crosan, R. T. A. (1999). The disjunction effect and reason based choice in games. *Organizational Behavior and Human Decision Processes*, 80, 118–133.
- Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology*, 41, 260–274.
- Dirac, P. A. M. (1958). *The principles of quantum mechanics*. Oxford, UK: Oxford University Press.
- Eisert, J., Wilkens, M., & Lewenstein, M. (1999). Quantum games and quantum strategies. *Physical Review Letters*, 83, 3077–3080.
- Falmagne, J. C., Regenwetter, M., & Grofman, B. (1997). A stochastic model for the evolution of preferences. In A. A. J. Marley (Ed.), *Choice, decision, and measurement: Essays in honor of R. Duncan Luce* (pp. 113–131). Mahwah, NJ: Erlbaum.
- Feynman, R. P., Leighton, R. B., & Sands, M. (1966). *The Feynman lectures on physics*, Vol. III. Reading MA: Addison Wesley.
- Fisher, D. L., & Goldstein, W. M. (1983). Stochastic PERT networks as models of cognition: Derivation of the mean, variance, and distribution of reaction time using order-of-processing (OP) diagrams. *Journal of Mathematical Psychology*, 27, 121–151.
- French, A. P., & Taylor, E. F. (1978). *An introduction to quantum physics*. NY: Norton.
- Gabora, L., & Aerts, D. (2002). Contextualizing concepts using a mathematical generalization of the Quantum formalism. *Journal of Experimental and Theoretical Artificial Intelligence*, 14, 327–358.
- Gardiner, C. W. (1991). *Quantum noise*. Berlin: Springer.
- Glimcher, P. W. (2005). Indeterminacy in brain and behavior. *Annual Review of Psychology*, 56, 25–56.
- Grossberg, S. (2000). The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Science*, 4, 233–246.
- Hameroff, S. R. (1994). Quantum coherence in microtubules: A neural basis for emergent consciousness. *Journal of Consciousness Studies*, 1(1), 91–118.
- Hughes, R. I. G. (1989). *The structure and interpretation of quantum mechanics*. Cambridge, MA: Harvard University Press.
- Kemeny, J. G., & Snell, J. L. (1960). *Finite Markov chains*. Princeton, NJ: Van Nostrand.
- Kolmogorov, A. N. (1933). *Foundations of the theory of probability*.
- Laming, D. R. (1968). *Information theory of choice-reaction times*. NY: Academic Press.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–105.
- Meyer, D. E., Yantis, S., Osman, A. M., & Smith, K. J. (1985). Temporal properties of human information processing: Tests of discrete versus continuous models. *Cognitive Psychology*, 17(4), 445–518.
- Miranker, W. L. (2002). A quantum state model of consciousness. *Journal of Consciousness Studies*, 9(3), 3–14.
- Lambert-Mogiliansky, A. L., Zamir, S., & Zwirn, H. (2004). Type indeterminacy: A model of the KT (Kahneman–Tversky)—Man. Paper presented at *Foundations Utility Risk and Decision Theory XI*, Paris, France.
- Nielsen, M. A., & Chuang, I. L. (2000). *Quantum computation and quantum information*. Cambridge, UK: Cambridge University Press.
- Penrose, R. (1989). *The emperor's new mind*. Oxford, UK: Oxford University Press.
- Pike, A. R. (1966). Stochastic models of choice behavior: Response probabilities and latencies of finite Markov chain systems. *British Journal of Mathematical and Statistical Psychology*, 19, 15–32.
- Piotrowski, E. W., & Sladkowski, J. (2003). An invitation to quantum game theory. *International Journal of Theoretical Physics*, 42, 1089.
- Pribram, K. H. (1991). *Brain and perception: Holonomy and structure in figural processing*. NJ: Erlbaum.
- Pribram, K. H. (1993). *Rethinking neural networks: Quantum fields and biological data*. Hillsdale, NJ: Erlbaum.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling accumulation of partial information. *Psychological Review*, 95, 238–255.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Regenwetter, M., Falmagne, J.-C., & Grofman, B. (1999). A stochastic model of preference change and its application to 1992 presidential election panel data. *Psychological Review*, 106, 362–384.
- Rumelhart, D., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 1. Cambridge, MA: MIT Press.
- Savage, L. J. (1954). *The foundations of statistics*. NY: Wiley.
- Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology*, 24, 449–474.
- Shankar, R. (1994). *Principles of quantum mechanics*. NY: Plenum Press.
- Shiffrin, R., & Maynard, T. (1988). Moments of transition—Additive random variables defined on finite, regenerative random processes. *Journal of Mathematical Psychology*, 32, 313–340.

- Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, 44, 408–463.
- Stapp, H. (1999). Attention, intention and will in quantum physics. *Journal of Consciousness Studies*, 6(8/9), 143–164.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.
- Townsend, J. T., & Busemeyer, J. R. (1995). Dynamic Representation of Decision Making. In R. Port, & T. van. Gelders (Eds.), *Mind As Motion: Explorations in the dynamics of cognition* (pp. 101–120). Cambridge, MA: MIT Press.
- Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science*, 3, 305–309.
- Vickers, D. (1979). *Decision processes in visual perception*. NY: Academic Press.
- Von Neumann, J. (1955). *Mathematical foundations of quantum mechanics*. Princeton, NJ: Princeton University Press.
- Woolf, N. J., & Hameroff, S. R. (2001). A quantum approach to visual consciousness. *Trends in Cognitive Science*, 5, 472–478.