J. R. Busemeyer

Psychological and Brain Sciences, Indiana University, Bloomington, In, 47405, USA
E-mail: jbusemey@indiana.edu

J. S. Trueblood

Cognitive Science Program, Indiana University, Bloomington, Indiana, 47405 USA
E-mail: jstruebl@indiana.edu

This chapter examines the empirical evidence for interference effects in psychological experiments. It also reviews the competing interpretations of these effects with respect to traditional cognitive models and new quantum cognition models.

Human behavior is not deterministic. If the same person is asked the same question on two different occasions, then there is some reasonable chance that the answers are inconsistent, even when there is no *known* intervention to account for this change. Behavior seems to vary across people, and within a person, it varies across occasions in ways that are far from perfectly predictable. Consequently psychologists need to use some type of probability theory to account for this indeterministic behavior. What kind of probability theory is best for modeling human behavior?

Almost all previous theoretical work in psychology has been developed along the lines of Kolmogorov's theory.[1] According to this theory, probabilities are assigned to events represented as sets from a sample space that form a Boolean algebra. While this has certainly proved to be a useful theory, the Boolean logic that lies at its foundation may be too restrictive to fully account for human behavior. Another probability theory is based on the intuitionistic logic of Brouwer.[2] According to this theory, probabilities are assigned to events represented by open sets from a topology. This theory relaxes the complementation axiom of Boolean algebra, and provides one way to generalize Kolmogorov's theory. A third probability theory is based on quantum principles.[3] According to this theory, probabilities are assigned to events represented as subspaces of a vector space. This theory relaxes the distributive axiom of Boolean algebra, which provides another way to generalize Kolmogorov's theory. This chapter explores the latter approach.

Why consider quantum probability for human behavior? One good reason is the pervasive finding of interference effects in psychology.[4] An interference effect is a violation of the law of total probability, which is a theorem of Kolmogorov theory derived from the distributive axiom. In other words, interference effects suggest that the distributive axiom may be violated in psychology. The finding of interference effects in particle physics was the primary reason for the construction of quantum theory by physicists from the beginning. What is an interference effect, what is the empirical evidence for these effects, what is the best explanation for these effects?

The purpose of this chapter is to answer these three questions, but before we do, we should note that there are two different lines of research using quantum theory in psychology. One is to develop a quantum physical model of the brain,[5] and the other is to develop models that are called 'quantum like'[4] or generalized quantum[6] or quantum structural[7] models. The latter are not quantum physics models of the brain, but instead they are mathematical models of human behavior derived from principles abstracted and extrapolated from quantum theory. This article is only concerned with the latter type of theory.

0.1. What is an interference effect?

Suppose we have two different judgment tasks: task A with J different levels of a response variable (e.g. $J = 2$ binary forced choice); and task B with K levels of a response measure (e.g. $K = 7$ point confidence rating). Participants are randomly assigned to two groups: group A receives only task A, but group BA receives task B followed immediately by task A. (Other variations are of course possible). We obtain estimates of the response probabilities for (a) $p(A = j) :=$ the probability of choosing level j from the response to task A from group A-alone, (b) $p(B = k) :=$ the probability of first responding with level k from task B, and (c) $p(A = j|B = k) :=$ the probability of responding with level j from task A given that the person responded with level k on the earlier task B. From the latter two probability distributions we can compute the *total probability* for response to task A as

$$p_T(A = j) = \sum_{k=1}^{K} p(B = k) \cdot p(A = j|B = k).$$

The interference effect for level j of the response to task A (produced by responding to task B) equals by definition

$$\delta_A(j) = p(A = j) - p_T(A = j).$$

Note that $\sum_{j=1}^{J} p(A=j) = 1 = \sum_{j=1}^{J} p_T(A=j)$ so that $\sum_j \delta_A(j) = 0$. Given these definitions, we can write

$$p(A=j) = p_T(A=j) + \delta_A(j).$$

Do these interference effects, $\delta_A(j)$, occur in human psychology experiments? If they do, how do we explain them?

0.2. What is the evidence for interference effects?

Below we summarize several lines of research that provide evidence for interference effects. The first three lines described below provide direct tests for interference effects, and the remaining three lines provide indirect tests of interference.

0.2.1. *Perception of ambiguous figures*

Interference effects were first investigated in the perceptual domain by Elio Conte.[8] Approximately 100 students randomly were divided into two groups: One was given 3 seconds to make a single binary choice (plus vs. minus) concerning an ambiguous figure A, and the other group was given 3 seconds to make a single binary choice for an ambiguous figure B followed 800 msec later by a 3 second presentation requesting another single binary choice (plus vs. minus) for figure A. The results produced significant interference effects. For example, for one type of testing stimuli, when test B preceded test A, the following results were obtained ($p(B+) :=$ probability of plus to figure B, $p(A+|B+) :=$ probability of plus to figure A given plus to figure B, $p_T(A+) :=$ total probability of plus to figure A, $p(A+) :=$ probability of plus to figure A alone):

$p(B+)$	$p(A+\|B+)$	$p(A+\|B-)$	$p_T(A+)$	$p(A+)$
.62	.78	.54	.69	.55

The interference effect equals $\delta_A(+) = p(A=+) - p_T(A=+) = .55 - .69 = -.14$, and $\delta_A(-) = +.14$.

0.2.2. *Categorization - decision making*

Townsend[9] and later Wang[10] use a paradigm to study the interactions between categorization and decision making, which is highly suitable for investigating interference effects. On each trial, participants are shown pictures of faces, which vary along two dimensions (face width and lip thickness). Two different distribution of faces were used: on average a 'narrow' face distribution had a narrow width and thick lips; on average a 'wide' face distribution had a wide width and thin lips. The participants were asked to categorize the faces as belonging to either a 'good' guy (G) or 'bad' guy (B) group, and/or they were asked to decide whether to take a 'attack' (A) or 'withdraw' (W) action. The participants were informed that 'narrow' faces had a .60 probability to come from the 'bad guy' (B) population, and 'wide' faces had a .60 chance to come from the 'good guy' (G) population. The participants were usually (.70 chance) rewarded for attacking 'bad guys' and they were usually (.70 chance) rewarded for withdrawing from 'good guys.' The primary manipulation was produced by using the following two test conditions, presented across a series of trials, to each participant. In the C-then-D condition, participants made a categorization followed by an action decision; in the D-Alone condition, participants only made an action decision.

The categorization-decision paradigm provides a simple test of the law of total probability. In particular, this paradigm allows one to compare the probability of taking an 'attack' (A) action obtained from the D-Alone condition with the total probability computed from the C-then-D condition. Townsend et al. (2000) reported significant deviations from the law of total probability using chi square tests, but they didn't examine the direction of these effects. The later study by Wang found a significant interference effect for the narrow faces and a smaller effect in the same direction for the wide faces. For example, using the narrow face data, when categorization preceded decisions, the following results were obtained ($p(G) :=$ probability categorize face as good guy, $p(A|G) :=$ probability attack given face categorized as good guy, $p(A|B) :=$ probability of attack given face categorized as a bad guy, $p_T(A) :=$ total probability to attack, $p(A) :=$ probability of attack when making only a decision):

$p(G)$	$p(A\|G)$	$p(A\|B)$	$p_T(A)$	$p(A)$
.17	.41	.63	.59	.69

Surprisingly, the probability of attacking without categorization was even higher than the probability of attacking after categorizing the face as a bad guy. The interference effect equals $\delta(A) = p(A) - p_T(A) = .69 - .59 = +.10$, and $\delta_A(W) = -.10$.

0.2.3. *Disjunction effect in decision making*

Perhaps the earliest report of interference effects is the disjunction effect.[11] The original studies were designed to test a rational axiom of decision theory called the sure thing principle.[12] According to the sure thing principle, if under state of the world X you prefer action A over B, and if under the complementary state of the world ˜X you also prefer action A over B, then you should prefer action A over B even when you do not know the state of the world. Tversky and Shafir experimentally tested this principle by presenting students with a two stage gamble, that is a gamble which can be played twice. At each stage the decision was whether or not to play a gamble that has an equal chance of winning $200 or losing $100 (the real amount won or lost was actually $2.00 and $1.00 respectively). The key result is based on the decision for the second play, after finishing the first play. The experiment included three conditions: one in which the students were informed that they already won the first gamble, a second condition in which they were informed that they lost the first gamble, and a third in which they didn't know the outcome of the first gamble. If they thought they won the first gamble, the majority (69%) chose to play again; if they thought they lost the first gamble, then again the majority (59%) chose to play again; but if they didn't know whether they won or lost, then the majority chose not to play (only 36% wanted to play again). Tversky and Shafir replicated this experiment using both a within subject design (the same person made choices under all conditions separated by a week) as well as with a between subject design (different groups of participants received known and unknown conditions).

This disjunction effect can be interpreted as an interference effect for the following reason. Define G as the event of playing the second gamble, W is observing a win on the first gamble, and L is observing a loss on the first gamble. The player can choose G or $\bar{G}$ alone (without observing the outcome of the first game); or the player can *observe* the outcome (W, L) first and then choose G or $\bar{G}$. Then $p(G)$ is the probability of gambling under the unknown condition, and $p_T(G) = p(W) \cdot p(G|W) + p(L) \cdot p(G|L)$ is the probability of choosing to gamble after observing the first play outcome. The total probability is a weighted average of the two known conditions, and so it requires that the probability of playing under the unknown condition must lie in between the two known probabilities. The results show that the probability for the unknown condition is below the smaller probability for the known condition. Therefore we have $p(G) < p(G|L) < p_T(G)$, which implies a negative interference effect.

This result is cited quite frequently, but the result remains controversial. Note that the gamble is actually quite attractive and it has a very positive expected value. Barkan and Busemeyer conducted a very similar study using the same gamble and under conditions in which participants chose to play the gamble a second time under three conditions: planning for a win or planning for a loss or without planning at all for the outcome of the first gamble; but the participants always preferred to play the gamble about 70% of the time and no disjunction effect was found.[13] Another study by Kuhberger et al. attempted a direct replication of Tversky and Shafir's gambling study, but they failed to find a disjunction effect:[14] if told that they won the first gamble, 60% chose to play again; if told that they lost the first gamble, 47% chose to play again; and if the first play was unknown, then 47% again choice to play again. There was no difference between the known loss condition and the unknown condition.

Another paradigm, using a prisoner dilemma game, was used by Shafir and Tversky[15] to test the sure thing principle. In all PDs, the two players need to decide independently whether to cooperate with the other player or to defect against the other. The player who stands to gain the most is the one who defects against a cooperating player. Mutual cooperation yields the second-highest payoff for each player. Mutual defection gives the players a payoff lower than that gained from mutual cooperation. Finally, the player who cooperates with a defective player gains the least. No matter what the other player does, an individual player always gains more when he defects; this makes defection the dominant option when the game is played only once against a given opponent (one-shot PD). A total of 80 participants were involved and each person played 6 PD games. Shafir and Tversky found that when a player was informed that the opponent defected, then 97% of the time they defected; if the player was informed that the other opponent cooperated, then 84% of the time they defected; but if they didn't know what the opponent chose then only 63% chose to defect.

Several other studies were conducted to replicate and extend the disjunction effect using the prisoner dilemma game. The first was done by Rachel Croson who used 80 participants, each playing 2 PD's, and half were required to predict or guess what the opponent would do and half were not asked to make this prediction.[16] In the first of Croson's experiments, the following results were obtained ($p(GD)$:= probability guess opponent defected, $p(D|GD)$:= probability player defects given opponent predicted to defect, $p(D|GC)$:= probability player defects given opponent predicted to cooperate, $p_T(D)$:= total probability to defect, $p(D)$:= probability to defect when opponent's action is unknown):

$p(GD)$	$p(D\|GD)$	$p(D\|GC)$	$p_T(D)$	$p(D)$
.54	.68	.17	.45	.23

The interference effect equals $\delta(D) = p(D) - p_T(D) = .23 - .45 = -.22$, and $\delta(C) = +.22$. The cooperation rates were much higher in this study as compared to the Shafir and Tversky (1992) study.

A later study by Li and Taplan also found evidence for disjunction effects but much weaker than Shafir and Tversky's original study.[17] Most recently, however, a very robust disjunction effect and replication of Shafir and Tversky (1992) was obtained by Matthew.[18] A total of 88 students played 6 PD games for real money against a computer agent. When told that the agent defected, then 92.4% defected; when told that the agent cooperated, then 83.6% defected; but when the agent's action was unknown only 64.9% defected.

0.3. What are explanations for these effects?

Interference effects are empirical results that need a scientific explanation. One cannot immediately jump to the conclusion that they are evidence for quantum mechanisms. Nor can one jump to the conclusion that interference effects are explained psychologically without quantum theory. The same 'psychological' explanation can be formulated probabilistically as either a Kolmogorov or a quantum model, and so it does not discriminate between these two theoretical competitors. The scientific way to determine which is best is to derive formal predictions from each theory and then compare the predictions with the data. The model that best predicts the experimental results is taken as the best explanation. Below we compare some competing explanations for these interference effects. We initially focus on the categorization - decision experiment, but the same models also apply to all of the findings summarized earlier.

0.3.1. *Markov model*

Markov models are commonly used in cognitive psychology. They provide the basis for random walk and diffusion models of decision making,[19] stochastic models of information processing,[20] and they are also the basis for multinomial processing tree models of memory retrieval[21] and memory recognition.[22]

Let us first consider a very simple Markov model for the categorization - decision making experiment proposed by Townsend et al. (2000). The person has to infer whether the face comes from the 'good' or 'bad' category (represented by two mutually exclusive Markov states $|B\rangle$ and $|G\rangle$, respectively), and given this inference, the person can intend to take an attack or withdraw action (represented by two mutually exclusive Markov states $|A\rangle$ and $|W\rangle$, respectively). The person starts in a state $|S\rangle$ determined by the face. Then $\phi(B|S)$ is the probability of transiting to inference state $|B\rangle$,and $\phi(G|S) = 1 - \phi(B|S)$ is the probability of starting in inference state $|G\rangle$. These probabilities form an initial distribution

$$\phi_0 = \begin{bmatrix} \phi(B|S) \\ \phi(G|S) \end{bmatrix}.$$

The distribution ϕ_0 deserves some additional comment and interpretation. According to the Markov model, at the beginning of a trial, the person enters exactly one of two states, the person either enters state $|B\rangle$ or enters state $|G\rangle$ and the person does not enter both states. The probabilities in ϕ_0 represent the theorist's uncertainty about the person's specific unknown state, and ϕ_0 is used to make predictions about the exact state at that moment.

If the player starts out inferring the face is a 'bad guy' $|B\rangle$, then the player can transit to the attack action state $|A\rangle$ with probability $\phi(A|B)$ or transit to the withdraw action state $|W\rangle$ with probability $\phi(W|B)$. If the player starts out inferring the face is a 'good guy' $|G\rangle$, then player can transit to the attack action $|A\rangle$ with probability $\phi(A|G)$ or transit to the withdraw action $|W\rangle$ with probability $\phi(W|G)$. These conditional probabilities form a 2×2 transition probability matrix

$$T = \begin{bmatrix} \phi(A|B) & \phi(A|G) \\ \phi(W|B) & \phi(W|G) \end{bmatrix}.$$

The sum across rows within each column must equal one for a transition matrix. This is required to guarantee that the probabilities sum to one after making a transition. After the transition from inference to action occurs, then the probability distribution across action states equals

$$\phi_1 = T \cdot \phi_0 \tag{0.1}$$
$$\begin{bmatrix} \phi(A|S) \\ \phi(W|S) \end{bmatrix} = \begin{bmatrix} \phi(A|B) \cdot \phi(B|S) + \phi(A|G) \cdot \phi(G|S) \\ \phi(W|B) \cdot \phi(B|S) + \phi(W|G) \cdot \phi(G|S) \end{bmatrix}.$$

This is the Chapman Kolmogorov equation for Markov models. The Chapman - Kolmogorov equation is simply a restatement of the law of total probability expressed in terms of the Markov states.

The probabilities in ϕ_1 deserve some further comment. As Equation 0.1 shows, there are two paths starting from the initial state $|S\rangle$ traveling to the final state $|A\rangle$. One is the path $|S\rangle \to |B\rangle \to |A\rangle$ that passes though the 'bad guy' inference, and the other is the path $|S\rangle \to |G\rangle \to |A\rangle$ that passes through the 'good guy' inference. The person can travel along one or the other but not both of these paths, because they are mutually exclusive. The person ends up either in the $|A\rangle$ state or the $|W\rangle$ state and not both. The particular action state that the person enters determines the choice response for an action. So the final probability distribution ϕ_1 represents the theorist's uncertainty about the person's final action state, and these probabilities are used by the theorist to predict the person's choices.

The preceding Markov model assumed that the states were directly observable. Now we explore the possibility that the states are mapped into responses by some 'noisy' process that allows measurement 'errors.' When measurements are noisy, it becomes important to introduce a distinction between states and observed responses. To do this, the categorization response is denoted by a variable C that can take on labels b or g for choosing the 'bad' or 'good' category respectively; and the choice response for an action is denoted by a variable D that can take on labels a or w for the choice of attack and withdraw actions, respectively.

The noisy measurement is achieved by introducing a probabilistic response map from states to responses. This is done by employing probabilistic state to response mappings. The following two matrices map inference states to category responses

$$C_b = \begin{bmatrix} C(b|B) & 0 \\ 0 & C(b|G) \end{bmatrix},\ C_g = \begin{bmatrix} C(g|B) & 0 \\ 0 & C(g|G) \end{bmatrix}.$$

For example, if the person enters the inference state $|B\rangle$, there is a probability $C(b|B)$ of categorizing the face as 'bad' and $C(g|B)$ that it is categorized as good. The next two matrices map action states to choices of each action

$$D_a = \begin{bmatrix} D(a|A) & 0 \\ 0 & D(a|W) \end{bmatrix},\ D_w = \begin{bmatrix} D(w|A) & 0 \\ 0 & D(w|W) \end{bmatrix}.$$

For example, if the person is in the state $|A\rangle$, then the person may actually choose to attack with probability $D(a|A)$, but the person may instead choose to withdraw with a probability $D(w|A)$. To guarantee that the categorization response probabilities sum to unity, we require that C_i contains probabilities such that $C_b + C_g = I$ where I is the identity matrix; to require the action response probabilities to sum to unity, we require that D_j contains probabilities such that $D_a + D_w = I$. This model reduces to the original Markov model without noise when we set $C(b|B) = 1 = C(g|G)$ and $D(a|A) = 1 = D(w|W)$. It is convenient to define a row vector $L = \begin{bmatrix} 1 & 1 \end{bmatrix}$ which used to sum across states.

Using these definitions, we can compute the following response probabilities. The probability that a face is categorized $C = i$ equals

$$p(C = i) = L \cdot C_i \cdot \phi_0.$$

If we observe the $C = i$ category response, then the probability distribution across inference states is revised by Bayes rule to become

$$\begin{aligned} \phi_i = \begin{bmatrix} \phi(B|C = i) \\ \phi(G|C = i) \end{bmatrix} &= \frac{1}{p(C = i)} \cdot C_i \cdot \phi_0 \\ &= \begin{bmatrix} \frac{\phi(B) \cdot C(i|B)}{p(C=i)} \\ \frac{\phi(G) \cdot C(i|G)}{p(C=i)} \end{bmatrix}. \end{aligned}$$

As the above equation shows, the categorization response changes the distribution across inference states. The probability of choosing action $D = j$ given that we observe a categorization response $C = i$ equals

$$p(D = j|C = i) = L \cdot D_j \cdot T \cdot \phi_i.$$

Therefore, the probability of choosing category $C = i$ and then choosing an action $D = j$ equals the matrix product

$$p(C = i, D = j) = L \cdot D_j \cdot T \cdot C_i \cdot \phi_0 \tag{0.2}$$

The probability that the face is first categorized as a 'bad guy' and then the person attacks equals

$$p(C = b, D = a) = L \cdot D_a \cdot T \cdot C_b \cdot \phi_0.$$

The probability that the face is first categorized as a 'good guy' and then the person attacks equals

$$p(C = g, D = a) = L \cdot D_a \cdot T \cdot C_g \cdot \phi_0.$$

The probability of attacking under the decision alone condition equals

$$\begin{aligned}
p(D = a) &= L \cdot D_a \cdot T \cdot \phi_0 \qquad (0.3)\\
&= L \cdot D_a \cdot T \cdot I \cdot \phi_0\\
&= L \cdot D_a \cdot T \cdot (C_b + C_g) \cdot \phi_0\\
&= L \cdot D_a \cdot T \cdot C_b \cdot \phi_0 + L \cdot D_a \cdot T \cdot C_g \cdot \phi_0\\
&= p(C = b, D = a) + p(C = g, D = a)\\
&= p_T(D = a).
\end{aligned}$$

In summary, this model satisfies the law of total probability, which of course fails to explain the interference effects found with the categorization decision making task.

This Markov model also applies to all of the other findings as follows. For the ambiguous figure results, we use states $|B\rangle$ and $|G\rangle$ to represent the plus or minus perceptions to figure B, and we use $|A\rangle$ and $|W\rangle$ to represent the plus or minus perceptions to figure A. For the two stage gambling game, we use states $|G\rangle$ and $|B\rangle$ to represent the 'win' or 'losses' inference about the first play of the gamble, and we use $|A\rangle$ and $|W\rangle$ to represent the 'play' or 'don't play' actions. For the PD game, we use states $|B\rangle$ and $|G\rangle$ to represent the 'defect' or 'cooperate' inference about the opponent, and we use $|A\rangle$ and $|W\rangle$ to represent the 'defect' or 'cooperate' strategy for the player. But this Markov model fails to explain any of the interference effects found in these other paradigms.

In fact, the matrix Equations 0.2 and 0.3 hold for any finite hidden Markov system. We could assume n inference states and m actions states for arbitrary n and m numbers of states. So these equations are not limited to a model with only two inference states and two action states with which we began. As long as the *same* initial state ϕ_0 and same transition matrix T is applied for both conditions (whether the inference state is measured or not), then the Markov model fails to account for the interference effects for all of these experiments.

0.3.2. *Quantum model*

The original explanation for the disjunction effect was a psychological explanation based on the failure of consequential reasoning under the unknown conditions. Shafir and Tversky (1992) explained the finding in terms of choice based on reasons as follows. Consider for example, the two stage gambling problem. If the person knew they won, then they had extra house money with which to play and for this reason they chose to play again; if the person knew they had lost, then they needed to recover their losses and for this other reason they chose to play again; but if they didn't know the outcome of the game, then these two reasons did not emerge into their minds. Why not? If the first play is unknown, it must definitely be either a win or a loss, and it can't be anything else. So the mystery is why these reasons don't emerge for the unknown condition. If choice is based on reasons, then the unknown condition has two good reasons. Somehow these two good reasons cancel out to produce no reasons at all! This sounds a lot like wave interference where one wave is rising and the other is falling. From this it follows that there is an interest in quantum models.

The psychological explanation given by Shafir and Tversky (1992) is quite consistent with a formal quantum mechanism for the effect. Busemeyer, Wang, and Townsend (2006) originally suggested a quantum interference interpretation for the disjunction effect, and since that time, various quantum models for this effect have been proposed, each one ultimately explaining the effects by interference terms, which includes Pothos and Busemeyer (2009), Khrennikov and Haven (2009), Aerts (2009), Yukalov and Sornette (2009) and Accardi, Khrennikov and Ohya (2009).

Busemeyer et al. (2006) started with the following simple quantum model.[1] Consider the category - decision making experiment once again. As in the Markov model, the person has to infer whether the face is a 'bad' or 'good' guy (represented by two mutually exclusive quantum states $|B\rangle$ and $|G\rangle$, respectively), and given this inference, the person may intend to attack or withdraw (represented by two mutually exclusive quantum states $|A\rangle$ and $|W\rangle$, respectively). Quantum theory replaces transition probabilities such as $\phi(A|B)$ with transition amplitudes such as $\langle A|B\rangle$ and $\phi(A|B) = |\langle A|B\rangle|^2$.

The person starts in a state $|S\rangle$. Then there is an amplitude $\langle B|S\rangle$ of transiting to the 'bad guy' inference and another amplitude $\langle G|S\rangle$ of transiting to the 'good guy' inference, $|\langle B|S\rangle|^2 + |\langle G|S\rangle|^2 = 1$. These amplitudes form an amplitude distribution (wave function) across inference states

$$\psi_0 = \begin{bmatrix} \langle B|S\rangle \\ \langle G|S\rangle \end{bmatrix}.$$

This amplitude distribution ψ_0 deserves more interpretation. If the face

[1] In quantum terminology, this model treats the inference as an observable operating within a two dimensional Hilbert space, and the action is another incompatible observable operating within the same Hilbert space.

is known to come from the 'bad guy' population, then $\langle B|S\rangle = 1$, and the initial state corresponds exactly to state $|B\rangle$; if the face is known to come from the 'good guy' population, then $\langle G|S\rangle = 1$, and the initial state corresponds exactly to state $|G\rangle$. But if $1 > |\langle B|S\rangle| > 0$ and $1 > |\langle G|S\rangle| > 0$, then the person is *not* exactly in state $|B\rangle$, and the person is *not* exactly in state $|G\rangle$ either. Furthermore, the person is *not* in both states at the same time. The person is exactly in an indefinite or superposition state represented by the wave function ψ_0. In the latter case, at a single moment in time, there is some *potential* to generate either one of the two mutually exclusive categorization responses. But only one of these potentials can become actualized to create an observed response.

On the one hand, if the state starts in the 'bad guy' inference, then there is an amplitude $\langle A|B\rangle$ of transiting to the attack action and another amplitude $\langle W|B\rangle$ of transiting to the withdraw action, $|\langle A|B\rangle|^2 + |\langle W|B\rangle|^2 = 1$. On the other hand, if the state starts in the 'good guy' inference, then there is an amplitude $\langle A|G\rangle$ of transiting to the attack action and another amplitude $\langle W|G\rangle$ of transiting to the withdraw action, $|\langle A|G\rangle|^2 + |\langle W|G\rangle|^2 = 1$. These transition amplitudes form a unitary matrix

$$U = \begin{bmatrix} \langle A|B\rangle & \langle A|G\rangle \\ \langle W|B\rangle & \langle W|G\rangle \end{bmatrix}.$$

Normally in quantum theory, the matrix U is require to be unitary: $U^\dagger U = I = UU^\dagger$. This is required to guarantee that the transformed amplitude distribution remains unit length, which is needed to guarantee that the final probabilities sum to one. The orthogonality restriction of the unitary matrix implies the equality

$$\langle A|B\rangle^* \langle A|G\rangle = -\langle W|B\rangle^* \langle W|G\rangle. \tag{0.4}$$

If we square the magnitudes of the entries of the unitary matrix, we obtain the transition probability matrix

$$T = \begin{bmatrix} |\langle A|B\rangle|^2 & |\langle A|G\rangle|^2 \\ |\langle W|B\rangle|^2 & |\langle W|G\rangle|^2 \end{bmatrix}$$

In order to satisfy the unitary property, this transition matrix must be doubly stochastic, that is both rows and columns sum to one. Double stochasticity implies that

$$\begin{aligned}|\langle W|B\rangle|^2 &= 1 - |\langle A|B\rangle|^2 = |\langle A|G\rangle|^2 \\ |\langle A|G\rangle|^2 &= |\langle W|B\rangle|^2 \\ |\langle A|B\rangle|^2 &= |\langle W|G\rangle|^2 \,.\end{aligned} \tag{0.5}$$

Equations 0.4 and 0.5 are a very strong constraints on this simple quantum model.

The final amplitude distribution across the action states is equal to

$$\begin{aligned}\psi_1 &= U \cdot \psi_0 \\ \begin{bmatrix}\langle A|S\rangle \\ \langle W|S\rangle\end{bmatrix} &= \begin{bmatrix}\langle A|B\rangle\langle B|S\rangle + \langle A|G\rangle\langle G|S\rangle \\ \langle W|B\rangle\langle B|S\rangle + \langle W|G\rangle\langle G|S\rangle\end{bmatrix}.\end{aligned}$$

The amplitude distribution ψ_1 across action states requires some comment. The amplitude $\langle A|S\rangle$ represents the direct path $|S\rangle \to |A\rangle$ from the initial state to the attack state. This path amplitude can be broken down by the theorist as the sum of the two other path amplitudes. One is the path $|S\rangle \to |B\rangle \to |A\rangle$ from the initial state to the 'bad guy' inference and then to the attack; and the other is the path $|S\rangle \to |G\rangle \to |A\rangle$ from the initial state to the 'good guy' inference and then to the attack. But we cannot conclude from this mathematical decomposition that the person passes through one or the other and not both of these two paths to get to the attack conclusion. Also one cannot conclude that the person travels both paths. Instead the person can travel directly from the initial state to the attack conclusion. Also the final amplitude distribution across action states is an indefinite or superposition state. One cannot conclude that the person ends up definitely in the $|A\rangle$ state or the $|W\rangle$ state and not both states immediately before the choice is made. Nor is the person exactly in both states. At that moment, both actions have some potential to be expressed, and the choice actualizes one of these potentials to produce the observed response.

If the person is exactly in the 'bad guy' state, then $\langle B|S\rangle = 1$ and the probability that the person attacks equals $|\langle A|B\rangle|^2$; if the person is exactly in the 'good guy' state, then $\langle G|S\rangle = 1$ and the probability that the person attacks equals $|\langle A|G\rangle|^2$. For the indefinite or superposed state, the action probabilities equal

$$
\begin{aligned}
|\langle A|S\rangle|^2 &= |\langle A|B\rangle\langle B|S\rangle + \langle A|G\rangle\langle G|S\rangle|^2 \\
&= \left|\langle A|B\rangle|^2|\langle B|S\rangle|^2 + |\langle A|G\rangle|^2|\langle G|S\rangle\right|^2 + \delta_1, \\
\delta_1 &= 2 \cdot \mathrm{Re}[\langle A|B\rangle\langle B|S\rangle\langle A|G\rangle\langle G|S\rangle] \\
|\langle W|S\rangle|^2 &= |\langle W|B\rangle\langle B|S\rangle + \langle W|G\rangle\langle G|S\rangle|^2 \\
&= \left|\langle W|B\rangle|^2|\langle B|S\rangle|^2 + |\langle W|G\rangle|^2|\langle G|S\rangle\right|^2 + \delta_2 \\
\delta_2 &= 2 \cdot \mathrm{Re}[\langle W|B\rangle\langle B|S\rangle\langle W|G\rangle\langle G|S\rangle].
\end{aligned}
$$

The probabilities from this quantum model can violate the law of total probability because of the cross product interference terms, δ_1 and δ_2, generated by squaring the sum. The orthogonality restriction from the unitary matrix (Eq. 0.4) implies that

$$
\begin{aligned}
\delta_1 &= 2 \cdot \mathrm{Re}[\langle A|B\rangle\langle A|G\rangle\langle B|S\rangle\langle G|S\rangle] \\
&= -2 \cdot \mathrm{Re}[\langle W|B\rangle\langle W|G\rangle\langle B|S\rangle\langle G|S\rangle] = -\delta_2.
\end{aligned}
$$

It is useful to express the complex numbers in complex exponential form:

$$
\begin{aligned}
\langle A|B\rangle^*\langle A|G\rangle &= |\langle A|B\rangle\langle A|G\rangle| \cdot e^{i\cdot\theta} \\
\langle W|B\rangle^*\langle W|G\rangle &= -\,|\langle A|B\rangle\langle A|G\rangle| \cdot e^{i\cdot\theta} \\
\langle B|S\rangle^*\langle G|S\rangle &= |\langle B|S\rangle\langle G|S\rangle| \cdot e^{i\omega}
\end{aligned}
$$

Then we obtain the well known formula for the quantum interference

$$
\delta_1 = 2 \cdot |\langle A|B\rangle\langle A|G\rangle\langle B|S\rangle\langle G|S\rangle| \cdot \cos(\theta + \omega) = -\delta_2.
$$

To account for the categorization - decision results obtained with the narrow faces, we can set $\langle B|S\rangle = \sqrt{.8}$ to approximate the observed initial probability of categorizing the narrow face as a bad guy (the actual value was .83 for the narrow faces), and we can set $\langle A|B\rangle = \sqrt{.60} = \langle W|G\rangle$, and $\langle W|B\rangle = \sqrt{.40}\cdot e^{-i\cdot 1.2313}$ and $\langle A|G\rangle = \sqrt{.40}\cdot e^{i\cdot 1.2313}$ to closely approximate the probabilities for actions conditioned on each categorization, while at the same time satisfying the requirements for a unitary matrix. Then the predicted probability of attacking under the decision alone condition equals $|\langle A|S\rangle|^2 = .69$, which closely approximates the results for the narrow face condition of the categorization - decision making experiment.

The problem that we run into when we apply this model to the PD game results of Shafir and Tversky is that the observed transition probabilities violate double stochasticity, and therefore they cannot be generated from a unitary matrix. For the PD game, define $C = b$ as the observation that the opponent has defected, $C = g$ as the observation that the opponent has cooperated, and $D = a$ as the defect response by the player. Then the results show that $p(D = a|C = b) = .97$ and $p(D = a|C = g) = .84$ but the unitary property requires the latter to be equal to $1 - p(D = a|C = b) = .03$ rather than .84. The unitary property is also violated by the two stage gambling game results of Tversky and Shafir. The perception results from Conte also violate the unitary property. The unitary property holds pretty well for the narrow face data obtained in the categorization - decision experiment by Wang, but it was violated by the wide face data. In short, the unitary property implied by the two dimensional model does not hold up well for this two dimensional model. To address this problem, Pothos and Busemeyer (2009) developed a four dimensional quantum model. But here we present a newer and much simpler two dimensional model.

0.3.3. *Quantum noise model*

On the one hand, the two dimensional Markov model failed because the interference effects violate the law of total probability. On the other hand, the two dimensional quantum model failed because the observed transition matrices violate double stochasticity. An interesting idea is to combine the two classes of models and form a quantum Markov model.[27] The following is a new model inspired by – but much simpler than – the Accardi et al. (2009) model. It assumes that the noisy measurements are used to assess the hidden quantum states.

Once again consider the analysis of the categorization - decision experiment. As before, $|S\rangle$ is the persons initial state, and we assume that there are two mutually exclusive states of inference: infer that the face belongs to the 'bad guy' category $|B\rangle$, or infer that the face belongs to the good guy category $|G\rangle$. The initial amplitude distribution is again represented by

$$\psi_0 = \begin{bmatrix} \langle B|S\rangle \\ \langle G|S\rangle \end{bmatrix}.$$

The initial state represents the amplitude distribution at the very beginning of the choice process, immediately after instructions. From these states the person can transition to two different intended actions $|A\rangle$ and $|W\rangle$ representing attack and withdraw. The initial amplitude distribution over inferences evolves for some period of time to produce a final amplitude distribution $\psi_1$over actions, which is used to make a choice. The final amplitude distribution is a unitary transformation of the initial distribution

$$\begin{aligned}\psi_1 &= \begin{bmatrix}\langle A|S\rangle \\ \langle W|S\rangle\end{bmatrix} \\ &= U \cdot \psi_0.\end{aligned}$$

The unitary transformation is defined as

$$\begin{aligned}U &= \begin{bmatrix}\langle A|B\rangle & \langle A|G\rangle \\ \langle W|B\rangle & \langle W|G\rangle\end{bmatrix} \\ &= \begin{bmatrix}\sqrt{u} & \sqrt{1-u}\cdot e^{i\cdot\theta} \\ -\sqrt{1-u}\cdot e^{-i\cdot\theta} & \sqrt{u}\end{bmatrix},\end{aligned}$$

for $0 \leq u \leq 1$, which satisfies the unitary property $U^{\dagger}U = UU^{\dagger} = I$ required to retain unit length following transformation.

According to this model, if the person starts out in state $|B\rangle$, then the person passes through one line of thought (with amplitude $\sqrt{u}$) that leads to one reason for attacking; if the person starts out in state $|W\rangle$, then the person passes through a different line of thought (with amplitude $\sqrt{1-u}$) that leads to a different reason for attacking; but if the person starts out in a superposition of these two states, then a direct path from $|S\rangle \rightarrow |A\rangle$ is taken in which the two lines of thought can constructively or destructively interfere (depending on the parameter θ).

The present model assumes that the quantum states are not directly observable because of measurement 'errors' or 'noise.' As before, it is important to distinguish between states and observed responses. Once again, the categorization response is denoted by a variable C that can take on labels b or g for choosing the 'bad' or 'good' category respectively; and the choice response for an action is denoted by a variable D that can take on labels a or w for the choice of attack and withdraw actions, respectively.

The choices are represented by measurement operators[28] that map states in observed responses. The two noisy measurement operators for categorizing as 'bad guy' or 'good guy' are defined by

$$C_b = \begin{bmatrix}\sqrt{C(b|B)} & 0 \\ 0 & \sqrt{C(b|G)}\end{bmatrix},\ C_g = \begin{bmatrix}\sqrt{C(g|B)} & 0 \\ 0 & \sqrt{C(g|G)}\end{bmatrix}.$$

For example, if the person is in the inference state $|B\rangle$, then the person may actually categorize the face as 'bad' with probability $C(b|B)$, but the person may instead categorize the face as good with probability $C(g|B)$. The two noisy measurement operators for choosing to attack or withdraw actions are defined by

$$D_a = \begin{bmatrix} \sqrt{D(a|A)} & 0 \\ 0 & \sqrt{D(a|W)} \end{bmatrix},\ D_w = \begin{bmatrix} \sqrt{D(w|A)} & 0 \\ 0 & \sqrt{D(w|W)} \end{bmatrix}.$$

For example, if the person is in action state $|A\rangle$, then the person may choose to attack with probability $C(a|A)$, but instead the person may choose to withdraw with probability $C(w|A)$. This model reduces to the original quantum model without noise when we set $C(b|B) = 1 = C(g|G)$ and $D(a|A) = 1 = D(w|W)$. These two measurement operators form a complete set because they satisfy the completeness property $C_b^\dagger C_b + C_g^\dagger M_g = I$ and $D_a^\dagger D_a + D_w^\dagger D_w = I$ needed to guarantee that the choice probabilities sum to one across actions.

Using this definitions, we can compute the following response probabilities. The probability that a face is categorized $C = i$ equals

$$p(C = i) = ||C_i \cdot \psi_0||^2.$$

If we observe the $C = i$ category response, then the amplitude distribution across inference states is revised by a quantum version of Bayes rule to become

$$\begin{aligned} \psi_i = \begin{bmatrix} \psi(B|C = i) \\ \psi(G|C = i) \end{bmatrix} &= \frac{1}{\sqrt{p(C = i)}} \cdot C_i \cdot \psi_0 \\ &= \begin{bmatrix} \frac{\psi(B)\cdot\sqrt{C(i|B)}}{\sqrt{p(C=i)}} \\ \frac{\psi(G)\cdot\sqrt{C(i|G)}}{\sqrt{p(C=i)}} \end{bmatrix}. \end{aligned}$$

As the above equation shows, the categorization response changes the amplitude distribution across inference states. The probability of choosing action $D = j$ given that we observe a categorization response $C = i$ equals

$$p(D = j|C = i) = ||D_j \cdot U \cdot \psi_i||^2.$$

Therefore, the probability of choosing category $C = i$ and then choosing an action $D = j$ equals the matrix product

$$p(C = i, D = j) = ||D_j \cdot U \cdot C_i \cdot \psi_0||^2 \tag{0.6}$$

The probability that the face is first categorized as a 'bad guy' and then the person attacks equals

$$p(C = b, D = a) = ||D_a \cdot U \cdot C_b \cdot \psi_0||^2.$$

The probability that the face is first categorized as a 'good guy' and then the person attacks equals

$$p(C = g, D = a) = ||D_a \cdot U \cdot C_g \cdot \psi_0||^2.$$

The probability of attacking under the decision alone condition equals

$$\begin{aligned} p(D = a) &= ||D_a \cdot U \cdot \psi_0||^2 \\ &= ||D_a \cdot U \cdot I \cdot \psi_0||^2 \\ &= ||D_a \cdot U \cdot (C_b + C_g) \cdot \psi_0||^2 \\ &= ||D_a \cdot U \cdot C_b \cdot \psi_0 + D_a \cdot U \cdot C_g \cdot \psi_0||^2 \\ &= p(C = b, D = a) + p(C = g, D = a) + \delta_A \\ &\neq p_T(D = a). \end{aligned} \tag{0.7}$$

In summary, this model can violate the law of total probability, and it can explain the interference effects found with the categorization decision making task.

To see how this works let us consider the two example applications where the original quantum model failed. First consider the two stage gambling game. In this case, we define $|B\rangle$:= inferring a loss on the first play, $|G\rangle$:= inferring a win on the first play, $|A\rangle$ choosing to play the second gamble, $|W\rangle$ choosing to not to play the second gamble, $C = g$ represents being told that you won the first round, $C = b$ represents being told that you lost the first round, and $D = a$ represents choosing to play the gamble again on the second round. Setting $C(b|B) = 1 = C(g|G)$, $D(a|A) = 1$, $D(a|W) = .28$, $u = .57$, and $\theta = .79 \cdot \pi$ exactly reproduces all of the results for the two stage gambling game reported by Tversky and Shafir (1992).

Next consider the results for the PD game. In this case, we define $|B\rangle$:= inferring opponent defects, $|G\rangle$:= inferring opponent cooperates, $|A\rangle$ player chooses to defect, $|W\rangle$ player chooses to cooperate, $C = g$ represents being told that opponent chose to cooperate, $C = b$ represents being told that opponent chose to defect, and $D = a$ represents player choosing to defect. Because the player is informed exactly about the opponent's decision, we simply set $C(b|B) = 1 = C(g|G)$. If we also set $D(a|A) = 1$, $D(a|W) = .68$, $u = .61$, and $\theta = \pi$, then this model produces the following results. if the opponent is known to defect, then the probability that the player defects equals $p(D = a|C = b) = .88$; if the opponent is known to cooperate, then the probability that the player defects $p(D = a|C = g) = .81$; and if the opponent's action is unknown, we set $\langle I_1|S\rangle = \langle I_2|S\rangle = \frac{1}{\sqrt{2}}$ to produce a probability of defection equals to

$p(A_1) = .69$. This approximates the results obtained in the PD game. This model can also exactly fit both the wide and narrow face data from the categorization - decision task as well as the results obtained with the ambiguous figures (details not shown here). In short, this model can perfectly fit many of the results demonstrating interference effects. But it does not provide a simple way to test double stochasticity. Also it has too many parameters relative to the number of data points produced by these experiments and so it is difficult to empirically test. New experiments are needed that generate more conditions and data points to test the model.

0.4. What next?

Quantum explanations for interference effects found in psychology have already made one important contribution. Quantum theory has provided a common way to understand a number of paradoxical findings that have never been connected before, nor even mentioned together in the same articles. By examining interference effects and providing a common quantum account of these effects, quantum theorists have organized a new and general and uniform way to think about all these seemingly unrelated problems. This is a step forward. We think that these initial promising steps made toward understanding all of the various interference effects are encouraging other researchers to begin examining quantum models in other applications in psychology.

What is needed next is stronger tests of these models. The applications reviewed above involve too many free parameters and too little data. For the simplest models – the two state Markov model and the two state quantum model – it was still possible to test the key properties (law of total probability and double stochasticity, respectively). Unfortunately, when these properties were tested, they failed for the simple models. The more complex models can account for the findings in a post hoc way, but they have not provide a strong empirical test with such small data sets. New experiments on interference effects are needed with many more conditions to provide more data points for testing these models.

Quantum probability is considered by many to be a very specialized probability theory that is only useful in Physics. We believe there are useful applications of this theory outside of Physics.[29] One of the founding fathers of quantum theory, Neils Bohr,[30] speculated on this possibility and so did David Bohm.[31] In fact, quantum and Markov probability theories are highly similar. Both Markov and quantum theories are defined by states and transition operators, and the only difference is the final way that the probabilities are computed. There are two key differences. Markov theory operates directly on probabilities and therefore it obeys the law of total probability but it does not have to obey the doubly stochastic law. Quantum theory operates on amplitudes and probabilities are obtained by squaring the amplitudes, consequently it does not have to obey the law of total probability but instead it must obey the law of double stochasticity.

References

1. A. N. Kolmogorov, *Foundations of the theory of probability* (N.Y.: Chelsea Publishing Co., 1933/1950).
2. L. Narens, *Theories of probability: An examination of logical and qualitative foundations.* (World Scientific Publishing Company, 2007).
3. J. Von Neumann, *Mathematical Foundations of Quantum Theory* (Princeton University Press, 1932/1955).
4. A. Y. Khrennikov, *Ubiquitous quantum structure: From Psychology to Finance* (Springer, 2010).
5. S. R. Hammeroff, *Philosophical Transactions Royal Society London (A)* **356**, 1869 (1998).
6. H. Atmanspacher and T. Filk, *Journal of Mathematical Psychology* **54**, 314 (2010).
7. D. Aerts, *Journal of Mathematical Psychology* **53 (5)**, 314 (2009).
8. E. Conte, A. Y. Khrennikov, O. Todarello, A. Federici, L. Mendolicchio and J. P. Zbilut, *Open Systems and Information Dynamics* **16**, 1 (2009).
9. J. T. Townsend, K. M. Silva, J. Spencer-Smith and M. Wenger, *Pragmatics and Cognition* **8**, 83 (2000).
10. J. R. Busemeyer, Z. Wang and A. Lambert-Mogiliansky, *Journal of Mathematical Psychology* **53 (5)**, 423 (2009).
11. A. Tversky and E. Shafir, *Psychological Science* **3**, 305 (1992).
12. L. J. Savage, *The foundations of statistics* (John Wiley & Sons, 1954).
13. R. Barkan and J. R. Busemeyer, *Psychological Bulletin and Review* **10**, 353 (1999).
14. A. Kuhberger, D. Komunska and J. Perner, *Organizational Behavior and Human Decision Processes* **85 (2)**, 250 (2001).
15. E. Shafir and A. Tversky, *Cognitive Psychology* **24**, 449 (1992).
16. R. Croson, *Organizational Behavior and Human Decision Processes* **80 (2)**, 118 (1999).
17. S. Li and J. Taplin, *Chinese Journal of Psychology* **44 (1)**, 25 (2002).
18. J. R. Busemeyer, M. Matthews and Z. Wang, *A Quantum Information Processing Explanation of Disjunction Effects.*, in *Proceedings of the 29th Annual Conference of the Cognitive Science Society and the 5th International Conference of Cognitive Science*, eds. R. Sun and N. Myake (Erlbaum, 2006), pp. 131–135.

19. R. Ratcliff and P. Smith, *Psychological Review* **111**, 333 (2004).
20. J. T. Townsend and G. F. Ashby, *Stochastic Modeling of Elementary Psychological Processes* (Cambridge University Press, 1983).
21. W. H. Batchelder and D. M. Reiffer, *Psychonomic Bulletin and Review* **6**, 57 (1999).
22. L. L. Jacoby, *Journal of Memory and Language* **30**, 513 (1991).
23. J. R. Busemeyer, Z. Wang and J. T. Townsend, *Journal of Mathematical Psychology* **50**, 220 (2006).
24. E. M. Pothos and J. R. Busemeyer, *Proceedings of the Royal Society, B.* **276 (1665)**, 2171 (2009).
25. A. Y. Khrennikov and E. Haven, *Journal of Mathematical Psychology* **53 (5)**, 378 (2009).
26. V. Yukalov and D. Sornette, *Theory and Decision* , 1 (2010).
27. L. Accardi, A. Y. Khrennikov and M. Ohya, *Open Systems and Information Dynamics* **16 (4)**, 371 (2009).
28. C. W. Gardiner, *Quantum noise* (Springer-Verlag, 1991).
29. J. R. Bruza, P. Busemeyer and L. Gabora, *Journal of Mathematical Psychology* **53 (5)**, 303 (2009).
30. N. Bohr, *Atomic theory and human knowledge* (Wiley, 1958).
31. D. Bohm, *Quantum Theory* (Dover, 1979).