# Theoretical and empirical reasons for considering the application of quantum probability theory to human cognition*

Jerome R. Busemeyer and Jennifer S. Trueblood
Indiana University, Bloomington Indiana

**Abstract**

This paper examines six theoretical reasons for considering the application of quantum theory to human cognition. It also presents six empirical examples that – while puzzling from a classic probability framework – are coherently explained by a quantum approach.

Quantum physics was created to explain puzzling findings that were impossible to understand using the older classical physical theory. In the process of creating quantum mechanics, physicists also created a new theory of probability. Can this new probability theory be usefully applied to fields outside of physics? This paper explores the application of quantum probability theory to the field of cognition and decision making. Almost all previous modeling in cognitive and decision sciences has relied on principles derived from classical probability theory. But these fields have also encountered puzzling findings that also seem impossible to understand within this limited framework. Quantum principles may provide some solutions. In the following paper, we first provide some general reasons for considering the application of quantum probability to human cognition[1], second we review findings that are very puzzling from a classic probability theory framework, finally we show how quantum theory provides a more coherent and unified account of these diverse findings.

## 1 Six reasons for a quantum approach to cognition

The first reason concerns the quantum concept of superposition. Classic cognitive models assume that at each moment a person is in a definite state with

*This material is based upon work supported by US National Science Foundation Grant No. 0817965. The authors thank the IU quantum group for comments and ideas including J. Michael Dunn, Andrew J. Hanson, Andrew Lumsdaine, Lawrence S. Moss, Gerardo Ortiz, Amr Sabry, and Jeremiah Willcock

[1] Sixe reasons are described but they can be boiled down to one (incompatability between subspaces used to define events). But we prefer to spell out the six implications of this one general reason.

respect to some judgment. Of course, it is not known what the person's true state is at each moment, and so the model can only assign a probability to a response with some value at each moment. But the model is stochastic only because it does not know exactly what trajectory (definite state at each time point) a person is following. In this sense, cognitive and decision sciences currently model the cognitive system as if it was a *particle* producing a definite sample path through a state space. Quantum theory works differently by allowing you to be in an *indefinite* state (formally called a *superposition* state) at each moment in time. Strictly speaking, being in an indefinite or superposition state means that the model *cannot* assume that you have a definite value with respect to some judgment scale at each moment in time. You can be in an indefinite state that allows all of these definite states to have *potential* (technically called probability amplitudes) for being expressed at *each* moment (Heisenberg, 1958). A superposition state provides a better representation of the *conflict, ambiguity, or uncertainty* that people experience at each moment. In this sense, quantum theory allows one to model the cognitive system as if it was a *wave* moving across time over the state space.

A second reason concerns sensitivity to measurement of the cognitive system. Traditional cognitive models assume that whatever we record at a particular moment reflects the state of the system as it existed immediately before we inquired about it. The answer to a judgment question simply reflects the state regarding this question just before we asked it. One of the more provocative lessons learned from quantum theory is that taking a measurement of a system creates rather than records a property of the system (Peres, 1998). Immediately before asking a question, a quantum system can be in an indefinite state. The answer we obtain from a quantum system is constructed from the interaction of the indefinite state and the question that we ask (Bohr, 1958). This interaction creates a definite state out of an indefinite state. We argue that the quantum principle of constructing a reality from an interaction between the person's indefinite state and the question being asked actually matches psychological intuition better for complex judgments than the assumption that the answer simply reflects a pre-existing state.

The third reason concerns the quantum concept of measurement incompatibility. The change in state that results after answering one question causes a person to respond differently to subsequent questions. Answering one question disturbs the answers to subsequent questions and the order of questioning becomes important. In other words, the first question sets up a context that changes the answer to the next question. Consequently, we cannot define a joint probability of answers simultaneously to a conjunction of questions, and instead we can only assign a probability to the sequence of answers. In quantum physics, order dependent measurements are said to be non-commutative and quantum theory was especially designed for these types of measures. Many of the mathematical properties of quantum theory arise from developing a probabilistic model for non-commutative measurements, including Heisenberg's (1927) famous uncertainty principle (Heisenberg, 1958). Question order effects are major concern for attitude researchers, who seek a theoretical understanding of these effects similar to that achieved in quantum theory (Feldman & Lynch, 1988).

The fourth reason is that human judgments do not always obey classic laws of logic and probability. The classic probability theory used in current cognitive and decision models is derived from the Kolmogorov axioms (Kolmogorov, 1933/1950), which assign probabilities to events defined as sets. Consequently, the family of sets in the Kolmogorov theory obey the Boolean axioms of logic, and one important axiom of Boolean logic is the distributive axiom. From this distributive axiom, one can derive the *law of total probability*, which provides the foundation for inferences with Bayes nets. However, as reviewed below, the law of total probability is violated by the results of many psychological experiments. Quantum probability theory is derived from the Dirac (Dirac, 1958) and von Neumann axioms (Von Neumann, 1932/1955). These axioms assign probabilities to events defined as subspaces of a vector space, and the logic of subspaces does not obey the distributive axiom of Boolean logic (Hughes, 1989).[2] The fact that quantum logic does not always obey the distributive axiom implies that the quantum model does not always obey the law of total probability (Khrennikov, 2010). Essentially, quantum logic is a generalization of classic logic and quantum probability is a generalized probability theory. Classic probability theory may be too restrictive to explain human judgments.

The fifth reason concerns an assumption of classic probability models called the principle of *unicity* (Griffiths, 2003).[3] A single sample space is proposed that provides a complete and exhaustive description of all events that can happen in an experiment.[4] We argue that it is over simplifying the extremely complex nature of our world. It becomes implausible to think that a person is able to assign joint probabilities to all different kinds of events. Quantum probability does not assume the principle of unicity (Griffiths, 2003), and this assumption is broken as soon as we allow incompatible questions into the theory which cause measurements to be non-commutative. Incompatible questions cannot be evaluated on the same basis, so that they require setting up different sample spaces. This provides more flexibility for assigning probabilities to events, and it does not require forming all possible joint probabilities, which is a property we believe is needed to understand the full complexity of human cognition and decision.

The last reason concerns the quantum concept of entanglement.[5] In cognitive science, judgments are often assumed to be “decomposable" so that the whole can be understood in terms of their constituent parts and how these are related together. This decomposability is reflected in the common assumption that there exists complete joint probability distribution across all of the ques-

[2] They do, however, form a partial Boolean algebra structure.

[3] This fifth reason is closely related to the third reason but spells out an important new consequence.

[4] Kolmogorov realized that different sample spaces are needed for different experiments, but his theory does not provide a coherent principle for relating these separate experiments. This is exactly what quantum probability theory is designed to do.

[5] The use of entanglement here refers only to the failure to construct a joint probability distribution for all pairwise events. It has nothing to do with the concept of reality or locality discussed in physics.

tions that can be asked, from which one can reconstruct any possible observed pairwise joint distribution for any pair of questions. Quantum probability theory allows systems to act in a non-decomposable manner such that pairwise probabilities cannot be derived from a common joint probability distribution. Intuitively this result suggests there is an extreme form of correlation between the systems which goes beyond the correlations derived from traditional probability theory.

Now that we have identified some general reasons for considering a quantum approach to cognition and decision, we review some puzzling empirical results that motivate the application of quantum probability to human cognition.

# 2 Six empirical examples from cognition and decision

## 2.1 Disjunction effect

The first example is a phenomenon discovered by Amos Tversky and Eldar Shafir called the disjunction effect (Tversky & Shafir, 1992). It was discovered in the process of testing a rational axiom of decision theory called the sure thing principle (Savage, 1954). 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 98 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 knew they won the first gamble, the majority (69%) chose to play again; if they knew 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 explained the finding in terms of choice based on reasons as follows. 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. Busemeyer, Wang, and Townsend (2006) originally suggested that this finding was an example of an interference effect similar to that found in the double slit type of experiments conducted in particle physics.

Consider the following analogy between the disjunction experiment and the classic double slit type of physics experiment. Both cases involve two possible paths: in the disjunction experiment, the two paths are inferring the outcome of either a win or a loss with the first gamble; for the double split experiment, the two paths are splitting the photon off into the upper or lower channel by a beam splitter. In both experiments, the path taken can be known (observed) or unknown (unobserved). Finally in both cases, under the unknown (unobserved) condition, the probability (of gambling for the disjunction experiment, of detection at a location for the two slit experiment) falls far below each of the probabilities for the known (observed) cases. So we speculate that for the disjunction experiment, under the unknown condition, instead of definitely being in the win or loss state, the student enters a superposition state that prevents finding a reason for choosing the gamble.

## 2.2 Categorization - decision interaction

The second example is based on a paradigm that provides an examination of the interaction between categorization and decision making (Townsend, Silva, Spencer-Smith, & Wenger, 2000), which we discovered is highly suitable for testing Markov and quantum models (Busemeyer, Wang, & Lambert-Mogiliansky, 2009). On each trial, participants were shown pictures of faces, which varied along two dimensions (face width and lip thickness). The participants were asked to categorize the faces as belonging to either a 'good' guy or 'bad guy' group, and/or they were asked to decide whether to take a 'attack' or 'withdraw' action. The primary manipulation was produced by using the following two test conditions, presented on different 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. In total, 26 undergraduate students from a Midwest university participated in the study, and each person participated for 6 blocks of C-D trials with 34 trials per block, and one block of D-alone trials with 34 trials per block. Little or no learning was observed (because instructions provided all of the necessary information) and so the C-D trials were pooled across blocks. The main results for the narrow face condition are shown in Table 1 below.

Table 1

| C-then-D | | | | | D-Alone |
|---|---|---|---|---|---|
| $p(G)$ | $p(A\|G)$ | $p(B)$ | $p(A\|B)$ | $p_T(A)$ | $p(A)$ |
| .17 | .41 | .83 | .63 | .59 | .69 |

The first and third columns indicate the probability of categorizing a face as a 'good guy' versus 'bad guy', and the second and fourth column indicate the probability of attacking conditioned on being a 'good guy' or a 'bad guy' respectively. The last column shows the probability of attacking under the D-alone condition. The column labeled $p_T(A)$ is the total probability, which is computed by the well known formula

$$p_T(A) = p(G) \cdot p(A|G) + p(B) \cdot p(A|B). \tag{1}$$

The critical prediction concerns the probability of attacking in the D-alone condition. According to the law of total probability, the attack probability in the D-alone condition should be a weighted average of the two probabilities of attacking conditioned on each categorization state. In other words, we should always find $p(A|G) \leq p(A) \leq p(A|B)$. However, when we look at the results in Table 1 above, we see that $p(A)$ substantially exceeds $p_T(A)$. More dramatic is the fact that the probability of attacking in the D-alone condition is even greater than the probability of attacking after categorizing the face as a 'bad guy' so that $p(A) > p(A|B)$.

## 2.3 Perception of ambiguous figures

A third example comes from a study in perception of ambiguous figures. Interference effects were first investigated in the perceptual domain by Elio Conte (Conte et al., 2009). 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.

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

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). The interference effect refers to the difference $p_T(A+) - p(A=+) = +.14$.

## 2.4 Conjunction and disjunction fallacies

A fourth example is an important probability judgment error, called the conjunction fallacy (Tversky & Kahneman, 1983). Judges are provided a brief story (e.g., a story about a woman named Linda, who used to be a philosophy student at a liberal university and who used to be active in an anti-nuclear movement). Then the judge is asked to rank the likelihood of the following events: event F (e.g., that Linda is now active in the feminist movement), event B (e.g., that Linda is now a bank teller), event 'F $\cap$ B' (e.g., Linda is currently active in the feminist movement and a bank teller). The conjunction fallacy occurs when option F $\cap$ B is judged to be more likely than option B (even though the latter contains the former) (Tversky & Kahneman, 1983). Students also tend to judge the disjunction F $\cup$ B to be less likely than the individual event F which is called the disjunction fallacy (Carlson & Yates, 1989).These effects are very robust and they have been found with many different types of stories. The conjunction fallacy is also obtained using betting procedures that do not involve even involve asking directly for probabilities (Sides, Osherson, Bonini, & Viale, 2002). The conjunction error is considered an interference effect for the following reason. Define $F$ as the event 'yes to feminism,' $B$ as the event 'bank teller,' and $S$ is the Linda story. According to the law of total probability $p_T(B|S) = p(F \cap B|S) + p(\overline{F} \cap B|S) > p(F \cap B|S)$, but the judgments produce $p(B|S) < p(F \cap B|S) < p_T(B|S)$ which implies a negative interference effect.

## 2.5 Overextension of Category membership

A fifth example arises from research on category membership. Hampton asked students to judge the strength of category membership for various natural items and found that they often rated the membership for a conjunction of two categories to be greater than one of the individual categories (Hampton, 1988b). For example when presented the item 'Cuckoo' they rated its strength (on a zero to one scale) for the category pet to be .575; and they rated its strength for the category bird to be 1.0; but they rated its strength for the category 'pet bird' to be .842. This is analogous to the conjunction fallacy described above. In a second study, Hampton found that students often rate the membership strength of an item to a disjunction of two categories to be smaller than the rating for one of the individual categories (Hampton, 1988a). For example, when presented with the item 'ashtray' they rated its strength (on a zero to one scale) for 'home furnishings' to be .7; they rated its strength for furniture to be .30; but they rated its strength for 'home furnishings or furniture' to be .25. This is analogous to the disjunction fallacy described above. These and other effects were later replicated by Hampton.

The overextension effect for the conjunction can be viewed as an interference effect by using the following interpretation. Define $p(A|x)$ as the probability that category $A$ is true given the item $x$, and $p(A \cap B|x)$ as the probability that category $A \cap B$ is true when given the item $x$.Then according to the law of total probability we have $p(A \cap B|x) < p_T(A|x)$ but the judgments demonstrate $p(A|x) < p(A \cap B|x) < p_T(A|x)$ which again implies a negative interference effect.

## 2.6 Memory recognition over-distribution effect

The final (sixth) example is found in research on memory recognition. The phenomenon of interest is observed in experiments that use a paradigm called the conjoint − recognition paradigm. Initially, participants are rehearsed on a set $T$ of memory targets (e.g., each member is a short description of an event). After a delay, a recognition test phase occurs, during which they are presented a series of test probes that consist of trained targets from T, related non-targets from a different set $R$ of distracting events (e.g. each member is a new event that has some meaningful relation to a target event), and unrelated set $U$ of non-target items (e.g. each member is completely unrelated to the targets). During the memory test phase, three different types of recognition instructions are employed: the first is a verbatim instruction (V) that requires one to accept only exact targets from $T$; the second is a gist instruction (G) that requires one to accept only distractors from the related non targets from $R$; the third is an instruction to accept verbatim or gist items (VorG), that is it requires one to accept probes from either from $T$ or $R$. Hereafter V represents the event 'accept as a target from $T$', G represents the event 'accept as a non target from $R$' and VorG represents the event 'accept as either a target from $T$ or a non target from $R$.' Note that $T \cap R = \varnothing$ , and so logically V and G are supposed to be mutually exclusive events. Also, logically the event VorG should equal the event V∪G, but this remains an empirical question.

Consider memory test trials that employ a test probe belonging to the target set $T$. If the verbatim question is asked, then probability of accepting the target is formally defined by the conditional probability $p(V|T)$; if the gist question is asked, then the probability of accepting the target is formally defined by the probability $p(G|T)$; finally if the verbatim or gist question is asked, then this is formally defined by the probability $p(V or G|T)$.

Logically, a probe $x$ comes from $T$ or $G$ but not both, implying that $p(V or G|T) = p(V|T) + p(G|T)$. The difference, $EOD(T) = p(V|T) + p(G|T) - p(V or G|T)$ is an episodic over distribution effect. A positive $EOD$ effect was obtained from 116 different experimental conditions (Brainerd & Reyna, 2008). All but 10% of the 116 studies produced this effect, and the mean value of the $EOD$ equals .18.

## 3 A unified explanation for interference effects

Applications of quantum theory to cognition and decision began over ten years ago with pioneering work by Aerts (Aerts & Aerts, 1994), Atmanspacher (Atmanspacher & Romer, 2002), Bordley (Bordley, 1998), Khrennikov (Khrennikov, 1999). Quantum probability theory provides a simple and coherent explanation for all six of the above empirical findings of interference effects. Various special versions of quantum theory have been developed for these findings, including models by Blutner (2009),Pothos and Busemeyer (2009), Khrennikov and Haven (2009), Franco (2009)(Franco, 2009), Aerts (2009), and Yukalov and Sornette (2008) and Busemeyer, Pothos, Franco, Trueblood (2011) (Busemeyer, Pothos, Franco, & Trueblood, 2011). But all of these models share the following simple ideas. Consider the disjunction effect described earlier for the two stage gamble. It is assumed the decision maker begins in a superposition state $|\psi\rangle$ which is a vector in an $N$ dimensional Hilbert space that represents the decision maker's potentials for different events. The event 'choose to play the second gamble' is represented by a projector $P_G$ which projects states onto the subspace consistent with this event. The event 'first play is a win' is represented by a projector $P_W$ that projects the state onto the subspace consistent with this fact. The event 'first play is a loss' is represented by a projector $P_L = I - P_W$ where $I$ is the identity operator. If the first play produces a win, then according to Lüder's rule the state that follows this observation equals

$$|\psi_W\rangle = \frac{P_W|\psi\rangle}{\|P_W|\psi\rangle\|}.$$

Given this new state, the probability to play again is

$$p\left(G|Win\right) = \|P_G|\psi_W\rangle\|^2 = \frac{\|P_G P_W|\psi\rangle\|^2}{\|P_W|\psi\rangle\|^2}.$$

Similarly, if the first play produces a loss, then the probability to play again is

$$p\left(G|Loss\right) = \frac{\|P_G P_L|\psi\rangle\|^2}{\|P_L|\psi\rangle\|^2}.$$

Now consider the probability of choosing to gamble on the second play when the outcome of the first play is unknown:

$$\begin{aligned} p\left(G|unknown\right) &= \|P_G|\psi\rangle\|^2 \\ &= \|P_G\left(P_W + P_L\right)|\psi\rangle\|^2 \\ &= \|P_G P_W|\psi\rangle + P_G P_L|\psi\rangle\|^2 \\ &= \|P_G P_W|\psi\rangle\|^2 + \|P_G P_L|\psi\rangle\|^2 + Int \\ Int &= \langle\psi|P_W P_G P_G P_L|\psi\rangle + \langle\psi|P_W P_G P_G P_L|\psi\rangle^* \end{aligned}$$

The term $Int$ is called the interference term, which can be positive, or negative, or zero. If it is zero, and there is no interference, and in this special cases$\|P_G|\psi\rangle\|^2$ obeys the law of total probability. But if the interference term is negative, then the law of total probability is be violated, which is consistent with the observed findings. The previous articles cited above provide explicit ways to represent this interference effect.

All six findings can be accounted by using the same model. For example, the categorization - decision making results are obtained by identifying $P_W$ with the event that the face is categorized as a 'good guy' and $P_L$ represents the categorizing the face as a bad guy, and $P_G$ represents the decision to attack. The perceptual experimental results are obtained by identifying $P_W$ with the event that the judgment for figure B is a plus, $P_L$ represents the case that the judgment for figure B is a minus, and $P_G$ represents the decision about figure A. As a final example, the conjunction fallacy results are explained by identifying $P_W$ with the event that 'Linda is a feminist' and $P_L$ represents the event 'Linda is not a feminist' and $P_G$ represents the event 'Linda is a bank teller.'

In summary, there are strong empirical reasons for calling into question classic probability theory as an adequate explanation for human judgment and decision making behavior. Quantum theory provides a simple account that unifies all of the diverse findings within a common theoretical framework. Finally, quantum theory introduces important new concepts, including superposition, compatibility, and entanglement, to help explain human cognition.

## References


Aerts, D. (2009). Quantum structure in cognition. *Journal of Mathematical Psychology*, *53 (5)*, 314-348.

Aerts, D., & Aerts, S. (1994). Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, *1*, 85-97.

Atmanspacher, H., & Romer, H., H. & Wallach. (2002). Weak quantum theory: complementarity and entanglement in physics and beyond. *Foundations of Physics*, *32*, 379-406.

Blutner, R. (2009). Concepts and bounded rationality: An application of Niestegge's approach to conditional quantum probabilities. In L. e. a. Acardi (Ed.), *Foundations of probability and physics-5* (Vol. 1101, p. 302-310).

Bohr, N. (1958). *Atomic physics and human knowledge.* New York, Wiley.

Bordley, R. (1998). Violations of compound probability principles: Toward a generalized Heisenberg uncertainty principle. *Operations Research*, *46*, 923–926.

Brainerd, C., & Reyna, V. F. (2008). Episodic over-distribution: A signature effect of familiarity without recognition. *Journal of Memory and Language*, *58*, 765-786.

Busemeyer, J. R., Pothos, E. M., Franco, R., & Trueblood, J. S. (2011). A quantum theoretical explanation for probability judgment errors. *Psychological Review*, *118 (2)*, 193-218.

Busemeyer, J. R., Wang, Z., & Lambert-Mogiliansky, A. (2009). Comparison of markov and quantum models of decision making. *Journal of Mathematical Psychology*, *53 (5)*, 423-433.

Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision making. *Journal of Mathematical Psychology*, *50*, 220-241.

Carlson, B. W., & Yates, J. F. (1989). Disjunction errors in qualitative likelihood judgment. *Organizational Behavior and Human Decision Processes*, *44*, 368-379.

Conte, E., Khrennikov, A. Y., Todarello, O., Federici, A., Mendolicchio, L., & Zbilut, J. P. (2009). Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Systems and Information Dynamics*, *16*, 1-17.

Dirac, P. A. M. (1958). *The principles of quantum mechanics.* Oxford University Press.

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73 (3)*, 421-435.

Franco, R. (2009). The conjunctive fallacy and interference effects. *Journal of Mathematical Psychology*, *53 (5)*, 415-422.

Griffiths, R. B. (2003). *Consistent quantum theory.* Cambridge University Press.

Hampton, J. A. (1988a). Disjunction of natural concepts. *Memory and Cognition*, *16*, 579-591.

Hampton, J. A. (1988b). Overextension of conjunctive concepts: evidence for a unitary model for concept typicality and class inclusion. *Journal of Experimental Psychology: Learning Memory and Cognition*, *14*, 12-32.

Heisenberg, W. (1958). *Physics and philosophy.* Harper and Row.

Hughes, R. I. G. (1989). *The structure and interpretation of quantum mechanics.* Harvard University Press.

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

Khrennikov, A. Y. (2010). *Ubiquitous quantum structure: From psychology to finance.* Springer.

Khrennikov, A. Y., & Haven, E. (2009). Quantum mechanics and violations of the sure thing principle: The use of probability interference and other concepts. *Journal of Mathematical Psychology*, *53 (5)*, 378-388.

Kolmogorov, A. N. (1933/1950). *Foundations of the theory of probability.* N.Y.: Chelsea Publishing Co.

Peres, A. (1998). *Quantum theory: Concepts and methods.* Kluwer Academic.

Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability model explanation for violations of 'rational' decision making. *Proceedings of the Royal Society, B.*, *276 (1665)*, 2171-2178.

Savage, L. J. (1954). *The foundations of statistics.* New York: John Wiley & Sons.

Sides, A., Osherson, D., Bonini, N., & Viale, R. (2002). On the reality of the conjunction fallacy. *Memory and Cognition*, *30*, 191-198.

Townsend, J. T., Silva, K. M., Spencer-Smith, J., & Wenger, M. (2000). Exploring the relations between categorization and decision making with regard to realistic face stimuli. *Pragmatics and Cognition*, *8*, 83-105.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjuctive fallacy in probability judgment. *Psychological Review*, *90*, 293-315.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science*, *3*, 305-309.

Von Neumann, J. (1932/1955). *Mathematical foundations of quantum theory.* Princeton University Press.

Yukalov, V. I., & Sornette, D. (2008). Quantum decision theory as quantum theory of measurement. *Physical Letters A*, *372*, 6867-6871.