



ELSEVIER

Contents lists available at ScienceDirect

## Journal of Mathematical Psychology

journal homepage: [www.elsevier.com/locate/jmp](http://www.elsevier.com/locate/jmp)

# Q1 Exploring the relations between Quantum-Like Bayesian Networks and decision-making tasks with regard to face stimuli

Q2 Catarina Moreira\*, Andreas Wichert

Instituto Superior Técnico, INESC-ID, Av. Professor Cavaco Silva, 2744-016 Porto Salvo, Portugal

## HIGHLIGHTS

- Representation of objects in an arbitrary  $n$ -dimensional vector space, enabling their comparison through similarity functions.
- Computation of quantum interference effects through vector similarity functions.
- Usage of contents of images to compute quantum interference parameters.
- Application of a Quantum-Like Bayesian Network to predict violations to the Sure Thing Principle.

## ARTICLE INFO

## Article history:

Available online xxxx

## Keywords:

Image categorization  
Quantum interference effects  
Multidimensional vector similarities  
Quantum-Like Bayesian Networks

## ABSTRACT

In this work, we propose to model the Categorization/Decision experiment from Busemeyer et al. (2009) with a Quantum-Like Bayesian Network. We also propose the representation of objects (or events) in an arbitrary  $n$ -dimensional vector space, enabling their comparison through similarity functions. The computed similarity value is used to set the quantum parameters in the Quantum-Like Bayesian Network model. Just like in the work of Pothos et al. (2013), we are not restricting our model to a vector in a two-dimensional space, but to an arbitrary multidimensional space.

In the end, we conclude that the vector representation of the contents of the images can explain the paradoxical findings and the violations of the laws of classical probability that were found in some works of the literature, suggesting that the contents of the images can already produce some quantum effects.

© 2016 Elsevier Inc. All rights reserved.

## 1. Introduction

Q3

The purpose of this work is to explore the applications of the formalisms of quantum mechanics to areas outside of physics, more specifically in domains regarding decision making and cognition.

Q4

Quantum cognition has emerged as a research field that aims to build cognitive models using the mathematical principles of quantum mechanics. In this sense, psychological (and cognitive) models benefit from the usage of quantum probability principles because they have many advantages over classical counterparts (Busemeyer, Wang, & Shiffrin, 2015). In quantum theory, events are represented as multidimensional vectors in a Hilbert space. This vector representation comprises potentially for the occurrence of all events at the same time. In quantum mechanics, this property refers to the superposition principle. Under a

psychological point of view, a quantum superposition can be related to the feeling of confusion, uncertainty or ambiguity (Busemeyer & Bruza, 2012). This vector representation neither obeys the distributive axiom of Boolean logic nor the law of total probability. It also enables the construction of more general models that can mathematically explain cognitive phenomena such as violations of the Sure Thing Principle (Khrennikov & Haven, 2009; Martínez-Martínez & Sánchez-Burillo, 2016), which is the focus of this study. Quantum probability principles have also been successfully applied in many different fields of the literature, namely in biology (Asano et al., 2012; Asano, Khrennikov, & Ohya, 2015), economics (Haven & Khrennikov, 2013; Khrennikov, 2009), perception (Conte, 2008; Conte et al., 2007), jury duty (Trueblood & Busemeyer, 2011), game theory (Brandenburger, 2010; Mura, 2005), order effects (Wang, Solloway, Shiffrin, & Busemeyer, 2014), opinion polls (Khrennikov & Basieva, 2014; Khrennikov, Basieva, Dzhafarov, & Busemeyer, 2014), etc.

Previously in the literature, Busemeyer, Wang, and Lambert-Mogiliansky (2009) studied the differences between a classical Markov and a quantum dynamical model in order to explain some violations of the law of classical probability theory in a

\* Corresponding author.

E-mail addresses: [catarina.p.moreira@ist.utl.pt](mailto:catarina.p.moreira@ist.utl.pt) (C. Moreira), [andreas.wichert@ist.utl.pt](mailto:andreas.wichert@ist.utl.pt) (A. Wichert).

<http://dx.doi.org/10.1016/j.jmp.2016.10.004>

0022-2496/© 2016 Elsevier Inc. All rights reserved.

**Categorization** experiment. Participants were presented with a set of digitally modified images of faces. Then, they had to first to categorize the face as *Good* or *Bad* and then perform the decision to either *Withdraw* or *Attack*. In the end, the proposed quantum dynamical model was able to accommodate the violations of the laws of classical probability theory by fitting the quantum parameters. This work demonstrated that quantum theory could be applied to build more general models to explain paradoxical situations found in cognitive psychology. More recently, more experiments to investigate the impact of quantum interference effects under the **Categorization** experiment have been performed in the work of Wang and Busemeyer (2016).

In the present work, we propose an alternative way to accommodate the paradoxical findings detected in the experiments of Busemeyer et al. (2009) and Townsend, Silva, Spencer-Smith, and Wenger (2000) that takes only into account the contents of the images and their vector similarities. The current work makes use of a Quantum-Like Bayesian model, initially introduced in the work of Moreira and Wichert (2014), and later developed in the works of Moreira and Wichert (2015a,b, 2016). The similarity is used to fit quantum interference parameters in the Quantum-Like Bayesian Network model. The main advantage of the proposed Quantum-Like Bayesian Network towards other cognitive models is its predictive nature and its scalability. By scalability we mean that the network structure of the proposed model is able to model more complex decision scenarios (scenarios that are modelled with several random variables). Moreover, through the representation of objects (or events) by their contents, one is able to perform vector similarities in an  $n$ -dimensional vector space and compute quantum parameters.

Approaching this categorization/decision experiment under a quantum probabilistic point of view is also important for several reasons (Pothos & Busemeyer, 2013). For instance, in the work of Pothos and Busemeyer (2009), the authors showed that a classical Markov model could not explain the violations to the Sure Thing Principle found in the experiment. Of course, one could always model a Markov model with extra hidden states and parameterizations to model these violations. However, this would lead to an exponential increase in complexity. Quantum probability theory is important for this reason. The geometric representation of events, which is present in quantum probability, does not exist in a classical setting. The main advantage of this geometrical representation is the ability of allowing the rotation from one basis into another in order to contextualize events and interpret events, providing great flexibility to decision-making systems.

## 2. Overview of probabilistic graphical models

In this section, we introduce the concepts of classical and Quantum-Like Bayesian Networks.

### 2.1. Classical Bayesian Networks

A classical Bayesian Network can be defined by a directed acyclic graph structure in which each node represents a different random variable from a specific domain and each edge represents a direct influence from the source node to the target node. The graph can represent independence relationships between variables, and each node is associated with a conditional probability table that specifies a distribution over the values of a node given each possible joint assignment of values of its parents (Koller & Friedman, 2009).

The full joint distribution (Russel & Norvig, 2010) of a Bayesian Network, where  $X$  is the list of variables, is given by:

$$Pr(X_1, \dots, X_n) = \prod_{i=1}^n Pr(X_i | Parents(X_i)). \quad (1)$$

The formula for computing classical exact inferences on Bayesian Networks is based on the full joint distribution (Eq. (1)). Let  $e$  be the list of observed variables and let  $Y$  be the remaining unobserved variables in the network. For some query  $X$ , the inference is given by:

$$Pr(X|e) = \alpha Pr(X, e) = \alpha \left[ \sum_{y \in Y} Pr(X, e, y) \right] \quad (2)$$

$$\text{where } \alpha = \frac{1}{\sum_{x \in X} Pr(X = x, e)}.$$

The summation is over all possible  $y$ , i.e., all possible combinations of values of the unobserved variables  $y$ . The  $\alpha$  parameter corresponds to the normalization factor for the distribution  $Pr(X|e)$  (Russel & Norvig, 2010). This normalization factor comes from some assumptions that are made in Bayes rule.

### 2.2. Quantum-Like Bayesian Networks

A more recent work from Moreira and Wichert (2014) suggested defining the Quantum-Like Bayesian Network in the same manner as in the work of Tucci (1995), replacing real probability numbers by quantum probability amplitudes.

In this sense, the quantum counterpart of the full joint probability distribution corresponds to the application of Born's rule to Eq. (1). An interesting discussion about the foundations of Born's rule can be found in the article of Deutsch (1988).

$$Pr(X_1, \dots, X_n) = \left| \prod_{i=1}^n \psi_{(X_i | Parents(X_i))} \right|^2. \quad (3)$$

The general idea of a Quantum-Like Bayesian network is that, when performing probabilistic inference, the probability amplitude of each assignment of the network is propagated and influences the probabilities of the remaining nodes. In other words, every assignment of every node of the network is propagated until the node representing the query variable is reached. Note that, by taking multiple assignments and paths at the same time, these trails influence each other and produce interference effects.

The quantum counterpart of the Bayesian exact inference formula corresponds to the application of Born's rule to Eq. (2), leading to:

$$Pr(X|e) = \alpha \left| \sum_y \prod_{x=1}^N \psi_{(X_x | Parents(X_x), e, y)} \right|^2. \quad (4)$$

Expanding Eq. (4), it will lead to the quantum interference formula:

$$Pr(X|e) = \alpha \left( \sum_{i=1}^{|Y|} \left| \prod_x \psi_{(X_x | Parents(X_x), e, y=i)} \right|^2 + 2 \cdot \text{Interference} \right)$$

$$\text{Interference} = \sum_{i=1}^{|Y|-1} \sum_{j=i+1}^{|Y|} \left| \prod_x \psi_{(X_x | Parents(X_x), e, y=i)} \right| \cdot \left| \prod_x \psi_{(X_x | Parents(X_x), e, y=j)} \right| \cdot \cos(\theta_i - \theta_j). \quad (5)$$

In the end, we need to normalize the final scores that are computed to achieve a probability value, because we do not have the constraints of double stochasticity operators. In classical Bayesian inference, normalization of the inference scores is also necessary

due to assumptions made in Bayes rule. The normalization factor corresponds to  $\alpha$  in Eq. (5).

Note that, in Eq. (5), if one sets  $(\theta_i - \theta_j)$  to  $\pi/2$ , then  $\cos(\theta_i - \theta_j) = 0$ , which means that the quantum Bayesian Network collapses to its classical counterpart. That is, they can behave in a classical way if one sets the interference term to zero. Moreover, in Eq. (5), if the Bayesian Network has  $N$  binary random variables, we will end up with  $2^N$  free quantum  $\theta$  parameters. Approaches to tune those parameters under a Quantum-Like Bayesian Network approach are still an open research question.

In the next section, we will propose a vector representation that takes into account vector similarities in order to compute the quantum  $\theta$  parameters required to perform inferences in Quantum-Like Bayesian Networks.

### 3. A vector similarity model to extract quantum parameters

Over the current literature, quantum parameters must be assigned manually to obtain a prediction in order to accommodate fallacies. In this work, we attempt to extend the paradigm of Quantum-Like Bayesian Networks to be predictive by representing events (in this case, images) as  $n$ -dimensional vectors and use these vector similarities to find the quantum  $\theta$  parameters. This vector representation is similar to the approach proposed in the work of Pothos, Bussemeyer, and Trueblood (2013), where the authors represent a person's beliefs/actions in an  $n$ -dimensional vector space and the similarity between the vectors is measured by a projection operator, which corresponds to the computation of the squared length of the projected vector.

The reader might be thinking why we should expect that the  $\theta$  parameter computed from the similarities of vectors should correspond to the interference term in the Quantum-Like Bayesian Network. In the book of Bussemeyer and Bruza (2012), it is stated that the  $\theta$  parameter that arises in quantum interference effects corresponds to the phase of the angle of the inner product between the projectors of two random variables. They also state that the inner product provides a measure of similarity between two vectors (where each vector corresponds to a superposition of events). If the vectors are unit length, then the Cosine Similarity collapses to the inner product. Also, in the work of Trueblood, Pothos, and Bussemeyer (2014), the authors mention that similarity is understood to be a function of the distance between two concepts in a psychological space. Given all these relations, we can assume that the similarities computed between two vectors representing the images of faces can be used to set quantum interference parameters, since they are both computing the inner product between two random variables and, consequently, we can assume a mathematical equivalence between the  $\theta$  parameters computed from similarities and the quantum  $\theta$  parameters corresponding to the interference terms in the quantum Bayesian model.

#### 3.1. Using Cosine Similarity to determine quantum parameters

Cosine Similarity is a metric that measures the similarity between two  $n$ -dimensional vectors through the cosine that they share between them. It is a widely used metric in several research fields, specially in Information Retrieval (van Rijsbergen, 2004). Given two  $n$ -dimensional vectors  $A$  and  $B$ , the Cosine Similarity measure is given by:

$$\text{cosine\_sim}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^N A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}. \quad (6)$$

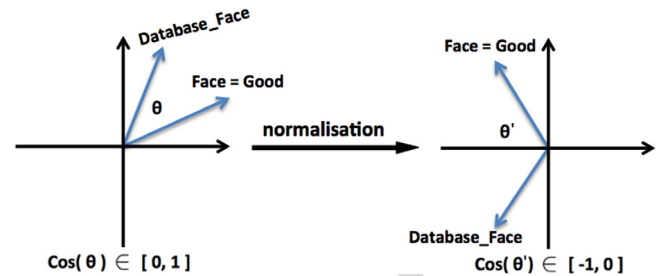


Fig. 1. Vector normalization to obtain quantum destructive interferences.

When mapping an image into an  $n$ -dimensional space, since the pixels of the images are always positive numbers between 0 and 1 (or between 0 and 255, depending on which scale), these vectors will always share a similarity angle:  $\cos(\theta) \in [0, 1]$ . This implies that when using this value in the quantum interference term in the Quantum-Like Bayesian Network inference formula (Eq. (5)), we will notice that the quantum probabilities will converge (or will be very close) to the classical probability. In the previous work of Yukalov and Sornette (2011), the authors noticed that, in order to accommodate the violations to the Sure Thing Principle, it was required that the quantum interference term be negative. In this sense, we applied a normalization to the vector representation of the images of the faces such that the new re-scaled vectors belong to the interval  $\cos(\theta') \in [-1, 0]$ . By doing this, we are re-scaling the vectors such that they cover the negative part of the vector space and to enable the occurrence of destructive quantum interferences. The re-scaling formula applied corresponds to the Min-Max Normalization formula, which is widely used in many different research fields, specially in Information Retrieval (van Rijsbergen, 2004). Note that the applied renormalization also enables constructive interferences. The renormalization spans the entire vector space producing both negative and positive interferences for each image representation. In this article, we focused on the destructive interferences, because to accommodate violations to the Sure Thing Principle these types of interferences (Yukalov & Sornette, 2010) are necessary.

$$Y_i = \text{MinMaxNorm}(X_i) = \frac{X_i - \min(X)}{\max(X) - \min(X)} \cdot (\text{new}_{\max} - \text{new}_{\min}) + \text{new}_{\min}. \quad (7)$$

Eq. (7) transforms a value  $X_i$  to  $Y_i$ , which fits in the range  $[\text{new}_{\min}, \text{new}_{\max}]$ . Fig. 1 illustrates the re-scaling process.

This vector representation of images (and events) also opens a new direction for the exploration of semantic similarities between concepts (Moreira & Wichert, 2015a,b).

Under the quantum mechanics point of view, quantum parameters that arise from interference effects represent the shift of energy waves. Under a quantum cognitive perspective, through this vector representation, they can be interpreted as correlations between events (beliefs) and the semantic relationships that they share between them.

In the next section, we present the experiment from Bussemeyer et al. (2009) and show how to apply the Quantum-Like Bayesian Network with the proposed vector similarity model in order to accommodate and predict violations to the Sure Thing Principle.

### 4. Empirical application of the Quantum-Like model to the categorization–decision experiment

In this section, we show how the Quantum-Like Bayesian Network model can be applied to predict the results obtained in the empirical experiments of Bussemeyer et al. (2009).





Fig. 2. Example of Wide faces used in the experiment of Busemeyer et al. (2009).

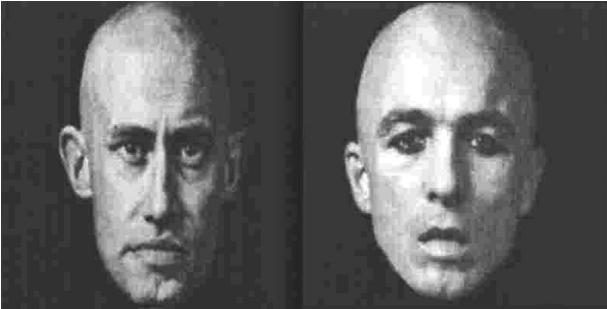


Fig. 3. Example of Narrow faces used in the experiment of Busemeyer et al. (2009).

#### 4.1. Categorization–decision making experiment

In the work of Busemeyer et al. (2009), the authors analysed the formalisms of quantum mechanics in order to describe the evolution of the cognitive process from the presentation of a decision problem to the actual decision. They performed an empirical experiment based on interactions between categorization and decision making. This experiment served as an empirical test to compare Classical Markov models with Quantum models. The experiment showed a violation of the law of probability theory while comparing the results between the probability of choosing a decision and the probability of making a categorization, followed by a decision. The authors proposed a Quantum Dynamical model that takes into account time evolution through the usage of Schrödinger's equation and unitary operators. A recent article from Yearsley and Pothos (2014) makes an interesting discussion about the classical notion of time under a quantum mechanical perspective (see Figs. 2 and 3).

The proposed experiment was the following. Given a set of images of faces, the participants had to categorize them as Good/Bad and had to make a decision towards that face: either to make an Attack or Withdraw. The Narrow faces had a 60% chance of belonging to the Lork group and 40% chance to the Adok group. The Wide faces, had a 40% chance of belonging to the Lork group and a 60% chance to the Adok group. Moreover, the Lork group is considered to be more hostile and therefore for 70% of them the right decision to take was to Attack. For the remaining 30%, the right decision was to Withdraw. For the Adok group, since they are considered more friendly, for 70% of the faces, the right decision was to Withdraw and for the remaining 30% was to Attack. Fig. 4 illustrates the distribution of the faces. The participants were divided into groups and had to perform four different tasks:

- One group had to perform first a categorization and then make a decision (the “C-then-D” condition);
- One group had to first make a decision and then perform a categorization (the “D-then-C” condition);
- One group had to just make the decision (“D-Alone” condition);

- One group had to just perform the categorization (“C-Alone” condition);

The main results obtained with this experiment are discriminated in Table 1. In this table,  $Pr(G)$  is the probability of a participant categorizing a face as Good.  $Pr(A|G)$  is the probability of a participant deciding to Attack, given that the face was categorized as being Good.  $Pr(B)$  is the probability of a participant categorizing a face as Bad.  $Pr(A|B)$  is the probability of choosing an Attack action, given that the face was categorized as Bad. Total Prob corresponds to the total probability through the formula  $Pr(A) Pr(A|G) + Pr(B)Pr(A|B)$ . Finally,  $Pr(A)$  corresponds to the total probability observed in the experiment. In order to verify if the experiment accommodates the law of total probability, the values obtained in the columns Total Prob and  $Pr(A)$  should be similar.

For the Wide faces, the classical law of total probability was not violated since the probability of choosing an Attack action alone is the same as the probability of Attack, but computed using the law of total probability formula. However, when we look at the results obtained with the Narrow faces, one can see that these probabilities are significantly different. When computing the probability of making an Attack with the law of total probability, the computed probability ended up in 59%. When computing the same probability is the D-Alone experiment, this probability increased to 69%. This deviation in the results suggests a violation to the Sure Thing Principle and leads to a violation of the law of total probability.

Next, we present an experimental simulation of the categorization/decision experiment performed by Busemeyer et al. (2009) using Quantum-Like Bayesian Networks and the proposed vector similarity model.

#### 4.2. Modelling the problem using Quantum-Like Bayesian Networks

The results observed in the empirical experiments of Busemeyer et al. (2009) can be represented in a Bayesian Network just like it is demonstrated on the left side of Fig. 5 for Narrow faces and the right side of the same figure for Wide faces. In the Bayesian Network both classical probabilities ( $Pr(X)$ ) and quantum probability amplitudes ( $\psi_x$ ) are specified.

#### 4.3. Computation of the probability of narrow faces

The first step to compute Bayesian inference consists in calculating the quantum version of the full joint probability distribution. This corresponds to Eq. (4). From the quantum-like full joint probability distribution, one can easily compute the probability of Attack in the following way (using Eq. (5)). Note that, in order to simplify the notation, we use the letter  $a$  instead of Attack, the letter  $g$  for Good, the letter  $b$  for Bad and the letter  $w$  for Withdraw. Also, the  $\alpha$  parameter is the normalization factor and corresponds to  $\alpha = [Pr(Attack) + Pr(Withdraw)]^{-1}$ :

$$Pr_{narrow}(Attack) = \alpha |\psi_{C=g}\psi_{D=a|C=g} + \psi_{C=b}\psi_{D=a|C=b}|^2 \quad (8)$$

$$Pr_{narrow}(Attack) = \alpha \left( |\psi_{C=g}\psi_{D=a|C=g}|^2 + |\psi_{C=b}\psi_{D=a|C=b}|^2 + Interf_A \right) \quad (9)$$

where,

$$Interf_A = 2 \cdot |\psi_{C=g}\psi_{D=a|C=g}| \cdot |\psi_{C=b}\psi_{D=a|C=b}| \cdot \cos(\theta_{a,g} - \theta_{a,b}).$$

In order to determine the normalization factor  $\alpha$ , one also needs to compute the probability  $Pr_{narrow}(Withdraw)$  in the same way:

$$Pr_{narrow}(Withdraw) = \alpha |\psi_{C=g}\psi_{D=w|C=g} + \psi_{C=b}\psi_{D=w|C=b}|^2 \quad (10)$$

$$Pr_{narrow}(Withdraw) = \alpha \left( |\psi_{C=g}\psi_{D=w|C=g}|^2 + |\psi_{C=b}\psi_{D=w|C=b}|^2 + Interf_W \right) \quad (11)$$



Fig. 4. Summary of the probability distribution of the Good/Bad faces in the experiment of Busemeyer et al. (2009).

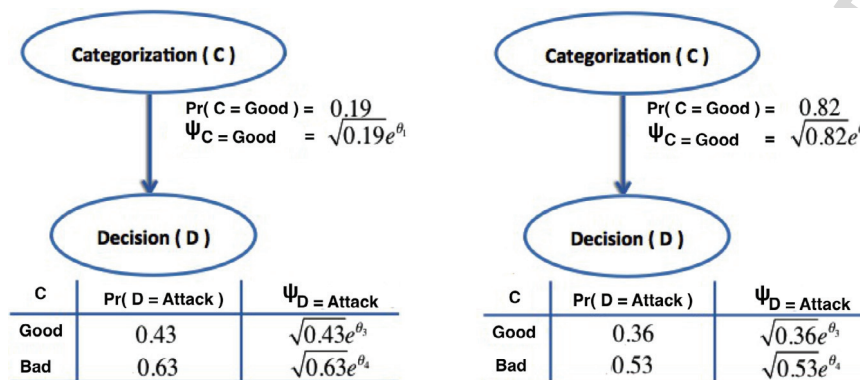


Fig. 5. Representation of the Narrow faces experiment (left) and Wide faces experiment (right) in a Bayesian Network with classical probabilities and quantum amplitudes. The classical probabilities are given by  $Pr(X)$  and the quantum amplitudes by  $\psi_x$ .

Table 1  
Empirical data collected in the experiment of Busemeyer et al. (2009).

C-then-D Experiments	$Pr(G)$	$Pr(A G)$	$Pr(B)$	$Pr(A B)$	Total prob	D-Alone $Pr(A)$
Empirical experiment (Busemeyer et al., 2009)						
Wide face	0.82	0.36	0.18	0.53	0.39	0.39
Narrow face	0.19	0.43	0.81	0.63	0.59	0.69

where,

$$Interf_W = 2 \cdot |\psi_{C=g} \psi_{D=w|C=g}| \cdot |\psi_{C=b} \psi_{D=w|C=b}| \cdot \cos(\theta_{w,g} - \theta_{w,b}).$$

The computation of the probabilities for the Wide faces is performed in an analogous way, but using the Quantum-Like Bayesian Network in Fig. 5.

In the next section, we present with more detail how quantum interference terms were computed using the images of the dataset in the experiments of Busemeyer et al. (2009) and the proposed similarity measure.

#### 4.4. Computing quantum interference terms

The quantum interference terms were obtained through vector similarities using a dataset of images of faces. The dataset used in our simulations is the same used in the experiments of Busemeyer et al. (2009), and it consists of 17 digitally modified Narrow faces (shape of face narrowed and lips enhanced) and 17 digitally modified Wide faces (shape of face widen and lips thick).

The images of the dataset used in the experiments of Busemeyer et al. (2009) were in greyscale. That is, they are represented by three matrices containing pixel information for the RGB colour scheme. In order to represent an image in an  $n$ -dimensional vector, it was required to convert the images into black and white to keep the main information of the image simple. This means that the image is represented by a single matrix in which the pixels are either 0 or 1. We made this conversion in order to obtain two different types of images: one that enhances the thickness of the lips and other features of the face (such as eyes and nose), and another one that diminishes the impact of these features. Fig. 6

shows an example of the conversion of the main dataset image into binary (black/white) images with the features either enhanced or reduced. The main motivation of doing this was to test if the content information of the image played any role in finding the quantum  $\theta$  parameters.

In the work of Busemeyer et al. (2009), the authors randomly considered a set of faces to be categorized as Good or as Bad according to some digitally modified features. In our work, since we did not have access to the information of which faces were classified as Good or as Bad, we decided to perform a simulation similar to the work of Busemeyer et al. (2009). For each simulation performed, we created several samples of this dataset, in which we randomly selected 70% of the Narrow faces to be considered Bad and 30% to be considered Good, just like it was already presented in Fig. 4. In the same way, we randomly selected 70% of the Wide faces to be Good and 30% to be Bad. For each image of the dataset, we converted the black/white face images into  $n$ -dimensional vectors. From this, we measured the similarity between every single image of the dataset with each of the faces that were classified as Bad and each of the images that were classified as Good. This would represent the categorization performed by each participant: given a face, he/she would have to categorize it as either Good or Bad. The similarity was computed through the Cosine Similarity function (Eq. (6)). The computed value was used to set the quantum interference terms in the Quantum-Like Bayesian Network and the final probability of the participant deciding to Attack was computed through Eq. (5). In the experiment performed by Busemeyer et al. (2009), this information was randomly assigned to each face of the dataset, and based on that information, they measured the probability of a participant answering the questions correctly. So, each face a priori already had a classification attached to it. We simulated this classification

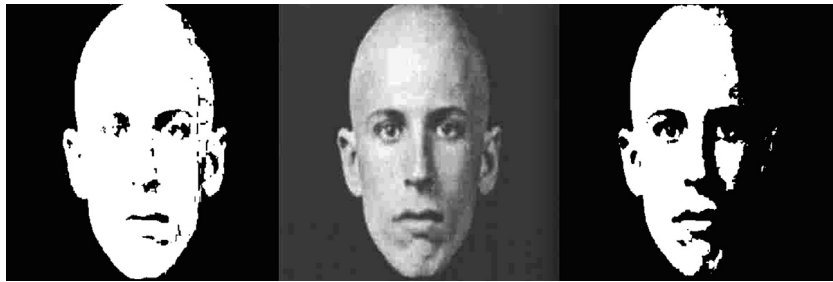


Fig. 6. Conversion of a dataset image into a binary image. Conversion with a small threshold (left). Conversion with a high threshold (right).

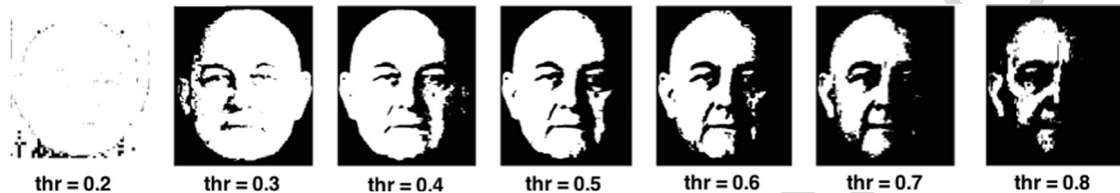


Fig. 7. Impact of the threshold when converting an image into a binary image. Threshold ranges from 0.2 (left) to 0.8 (right).

by randomizing the dataset 100 times and considering different faces to be either *Good* or *Bad* according to the percentages used in the original study. The mean value of the probabilities for each simulation was computed. Originally, it was this randomization that was the cause of the occurrence of the levels of uncertainty in the study of Busemeyer et al. (2009), leading to a violation of the laws of probability theory. In other words, one can see this uncertainty as a cause for the emergence of quantum interference effects. In the same way, we use the differences between the contents of the images, when compared to an image that was previously classified as either *Good* or *Bad* (since we do not know which classifications were attributed to each face), in order to raise uncertainty and to measure the quantum effects. Note that some readers might think that the quantum interference effects occur due to the fact that two images can be incompatible. However, one should take into account that incompatibility only means that you have to represent each of the questions (under a quantum-like point of view) in different basis, it does not mean that the two images are incompatible (in the experiment, people answered two incompatible questions about the same image).

#### 4.5. The impact of the conversion threshold

Also, in this experiment, we wanted to verify how the conversion of the images influences the probabilities of an *Attack* action under the Narrow faces. We varied the conversion threshold from [0.2, 0.8]. Fig. 7 shows an example of how the images vary according to the conversion threshold.

For each threshold, we analysed their respective probability distributions using histograms. We also fitted a normal probability distribution curve in order to check if the frequency of the occurrence of choosing the action *Attack* is distributed around the mean value. This will play an important role in choosing which is the best conversion threshold that describes the probability distribution of our data. Figs. 8–14 show the histograms of the experiments for each threshold with their respective normal density probability distributions.

Regarding this model, a legitimate question that one can pose is concerned with the assumption that the angles computed through vector similarities can be used as quantum interference terms. The answer to this concern can be addressed by taking into account how the dataset was built. In the experiment of Pothos and Busemeyer (2009), the authors performed digital

modifications in the dataset in order to enhance the properties that they wanted participants to perceive during the experimental setup: (1) enhance the narrowness or wideness of the shape of the faces and (2) enhance the lips making them more thick for narrow faces or thinner for wide faces. By doing these digital modifications in the dataset, the authors were already introducing some kind of categorization in the images of the faces: narrow faces tend to be part of the Lork group, which have a higher probability to be aggressive; and wide faces tend to be part of the Adok group, which are more friendly. By manipulating the images pixels, the authors were already encoding some kind of semantical/psychological information. By comparing the image vector information between a face categorized as *Good* with a face categorized as *Bad*, we are measuring the phase of the angle of the inner product between the projectors of two random variables, and this is precisely the definition of quantum interference term given in the book of Busemeyer and Bruza (2012). Moreover, the usage of vector similarities in order to represent quantum interferences has been already applied in previous literature, such as in the work of Pothos et al. (2013), where the author represents concepts in a multidimensional vector space and measures similarities between them.

From these histograms, one can observe that the probability distribution of Figs. 8 and 13 is a little skewed. That is, the highest occurrence of the probability of choosing the *Attack* action is shifted either to the left or to the right of the mean value of the distribution. On the other hand, the histograms that better describe the data by a normal distribution fit are Figs. 10–12. Note that, in the end, for each threshold image, the mean values are located between 0.64 and 0.65. This means that, the choice of the conversion threshold of the images does not have a significant impact on the final outcome of the results. However, there are some thresholds that lead to a probability distribution closer to a normal density probability function (which is the case of Figs. 10–12). For this reason, we chose the threshold that leads to the higher mean value of choosing an *Attack* action. In this case, Fig. 10, which corresponds to a conversion threshold of 0.4, that is, a threshold that slightly diminishes the features of the images. In Fig. 15, it is illustrated how the probabilities are distributed for the 100 samples tested using a conversion threshold of 0.4.

#### 4.6. Results and discussion

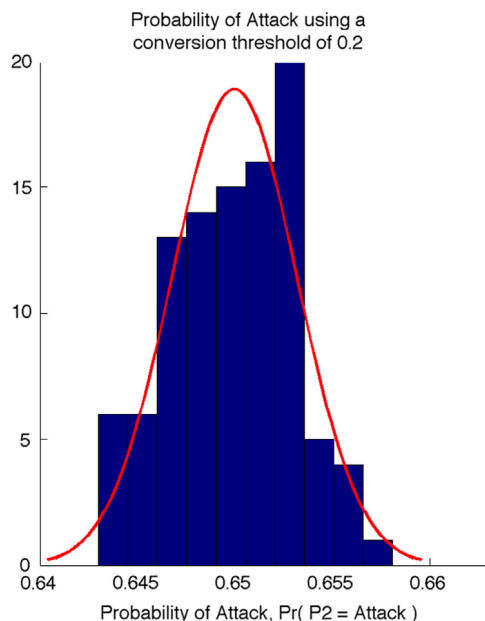
The results obtained after running the simulations described in the above sections are presented in Table 2. In the experiments



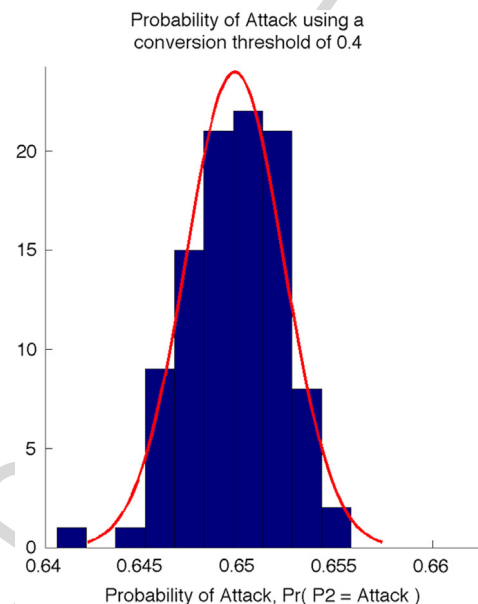
**Table 2**

Results from the application of the Quantum Like Bayesian Network (QLBN) model to the categorization/decision experiment and comparison with the Quantum Dynamical model (QDM) proposed in the work of Busemeyer et al. (2009).

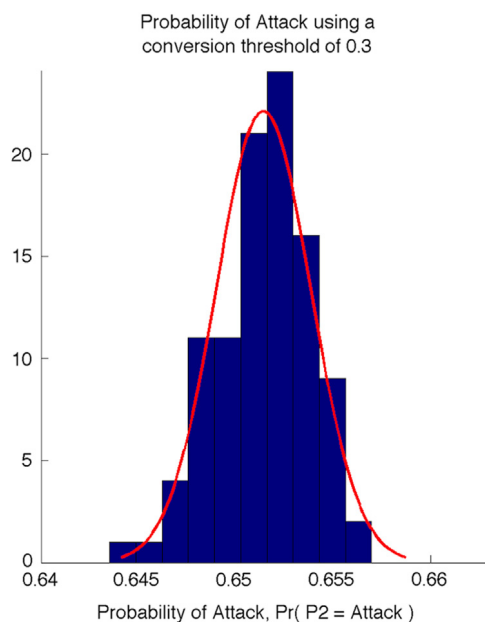
Literature	Pr(Attack) observed	Pr(Attack) classical probability	Pr(Attack) QDM	Error QDM	Pr(Attack) QLBN	Error QLBN
Narrow faces	0.69	0.59	0.74	5.00%	0.65	4.00%
Wide faces	0.39	0.39	0.39	0.00%	0.35	4.00%



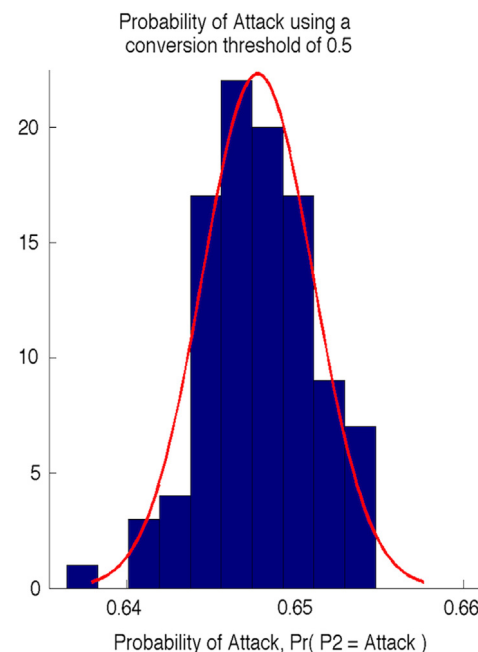
**Fig. 8.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.2.



**Fig. 10.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.4.



**Fig. 9.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.3.

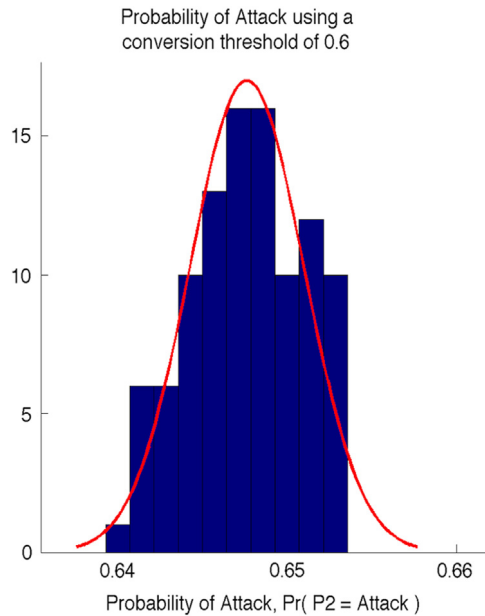


**Fig. 11.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.5.

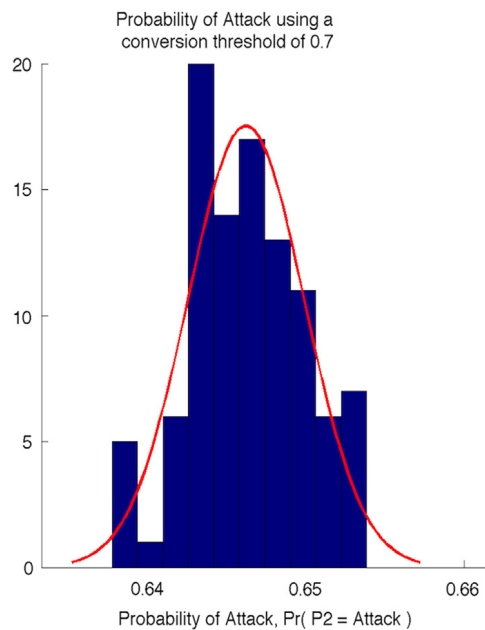
performed by Busemeyer et al. (2009), only the Narrow faces experiment presented a violation to the Sure Thing Principle. The Wide faces experiment presented the same results as the classical probability, so no violations occurred.

In the previous work of Busemeyer et al. (2009), the authors present a Quantum Dynamical Model (QDM) model to perform quantum time evolution. This model requires the creation of a doubly stochastic matrix, which represents the rotation of the

participants' beliefs (it can be favouring an *Attack* action or a *Withdraw* action). The double stochasticity is a requirement in order to preserve unit length operations and to obtain a probability value that does not require normalization. The time evolution of the model simulates the participants' deliberation process, until a final decision is reached, and is modelled using Schrödinger's Equation. To obtain the observed results depicted in Table 2, the



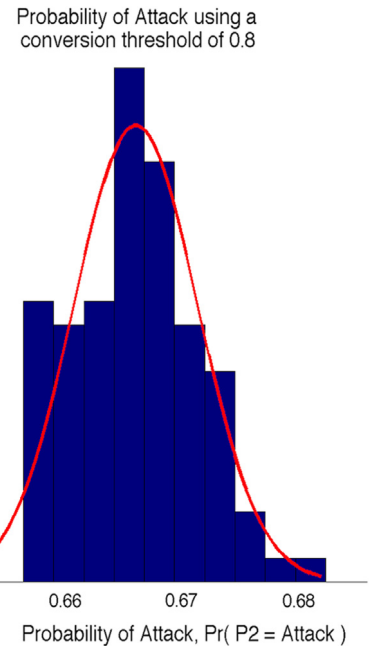
**Fig. 12.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.6.



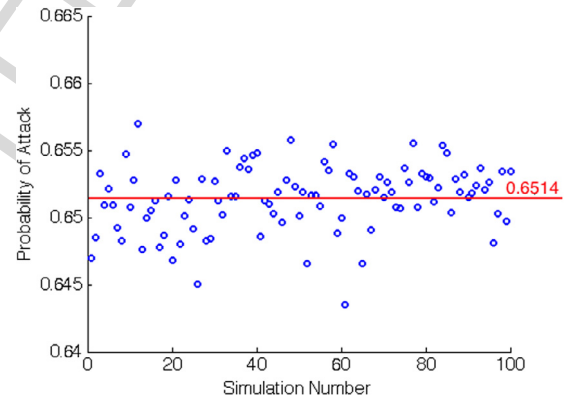
**Fig. 13.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.7.

authors had to fit the parameters of their model to this outcome. In the end, the Quantum Dynamical Model proposed in Bussemeyer et al. (2009) obtained an error percentage 5.00% for the Narrow faces and 0.00% for the Wide faces experiment. Note that the Quantum Dynamical Model fits three parameters in order to estimate four data points (the first four entries of Table 1).

With the proposed Quantum-Like Bayesian Network together with the geometric representation of events, we were able to build a quantum-like model that has a predictive nature. In a way, we also make use of quantum interference effects in order to explain the violations to the Sure Think Principle. The proposed model also has a predictive nature since the parameter fitting of the quantum model is found by geometric similarities. By predictive we mean that the model does not require any *a priori* knowledge about the



**Fig. 14.** Distribution of the probability of *Attack* through all simulations for a threshold of 0.8.



**Fig. 15.** Probability distribution of the 100 simulations performed when converting a grayscale image into a binary one with a threshold of 0.4.

outcome of the experiment in order to accommodate the violations to the Sure Thing Principle.

In the proposed model, the contents of the images (the pixels) are represented in an  $n$ -dimensional vector space. From this representation, we computed the geometric similarity between them through the usage of the cosine similarity measure. Since the contents of the images (pixels) are always positive (ranging between 0 and 1), it was required to **renormalize** this information in order to obtain quantum interference effects. Taking into account this **normalization** in the computation of the final quantum probabilities, one could predict the observed results with an error percentage of 4.00%. This preliminary result suggests that there could be a relation between quantum parameters with the semantic and geometric representation of events. We are aware that this is just a preliminary conclusion, and more experiments in this direction need to be conducted. In what concerns the Wide faces, no violations to the Sure Thing Principle were reported, since the probability of *Attack* in the *D-Along* condition was the same as the one computed using the law of total probability. In this case, the proposed model was able to predict the result with an error percentage of also 4.00%. In the end, the proposed similarity model tends to be more effective in decision scenarios that violate



the Sure Thing Principle. Moreover, the Quantum-Like Bayesian did not obtain very significant error rates when compared to the Quantum Dynamical Model. This means that the proposed model tends to have a similar performance when compared to state of the art models with the advantage of being able to estimate quantum interference parameters. This makes the model general, scalable and predictive.

Regarding scalability, it is well known that the computational costs of performing probabilistic inferences in Bayesian Networks grow with the number random variables (Koller & Friedman, 2009). That is why, for very complex decision scenarios, approximative methods are used in order to perform probabilistic inferences (Murphy, 2012). Bayesian Networks are decision support systems that are used very frequently to model complex decision scenarios such as Bioinformatics, in which many scenarios include dealing with an exponential number of genes (Zou & Conzen, 2005), medical decision support, where Bayesian networks are used to compute the probability of a patient having cancer given several conditions (Kahn, Roberts, Shaffer, & Haddawy, 1997), spam filters, in which the probability of some textual content being considered spam is computed (Sahami, Dumais, Heckerman, & Horvitz, 1998), etc. This is what we mean by generalization. Bayesian Networks are widely accepted structures in the literature, because they can deal with a big number of random variables and be applied in different decision scenarios.

## 5. Conclusion and final discussion

In this work, we propose to model the Categorization/Decision experiment from Busemeyer et al. (2009) with a Quantum-Like Bayesian Network. We also propose the representation of objects (or events) in an arbitrary  $n$ -dimensional vector space, enabling their comparison through similarity functions. The computed similarity value is used to set the quantum parameters in the Quantum-Like Bayesian Network model. Just like in the work of Pothos et al. (2013), we are not restricting our model to a vector in a multidimensional psychological space, but to an arbitrary multidimensional space.

The reason of choosing to address this problem through a Quantum-Like Bayesian Network approach, is because Bayesian Networks can model more complex decision scenarios very easily. They can generalize to more complex decision scenarios due to its network structure and compute probabilistic inferences more efficiently than, for instance, quantum projection-based models. In a quantum projection approach, this would be intractable with the increase of the number of random variables with the complexity of the decision scenario. However, in the quantum cognition literature, there is not much data available with more than two random variables. For this reason, we cannot verify the effectiveness of the proposed Bayesian network under more complex scenarios. The only thing we can do is to perform experiments with the available data.

An interested reader might also think how can the proposed Quantum-Like Bayesian Network be applied in other types of quantum cognition problems such as order effects. In social sciences and behaviour research, order effects are a kind of problem that consists in asking two consecutive questions in different order and obtaining a different answer for the same questions. That is, the context of the previous question influences the answer of the second question. Very generally speaking, we would say that there is a possibility to represent such problems due to the acyclic nature of the Bayesian Network. This means that, when performing probabilistic inferences, the information trails in the network tend to be directional, so this opens a door in modelling problems related to order effects. More research studies under this direction are necessary in order to verify this approach.

The results of the simulations of the experiment of Busemeyer et al. (2009) demonstrated that the proposed method is general and was able to reproduce the experimental observations of the violations of the Sure Thing principle with a small error percentage. We are aware that this is just a preliminary result and more experiments and studies are needed towards this direction in order to verify the applicability of this type of modelling in decision problems that are violating the Sure Thing Principle.

In the end, in this model we are assuming that the similarities computed between two vectors representing the images of faces can be used to set quantum interference parameters, since they are both computing the inner product between two random variables and, consequently, there is a mathematical equivalence between the  $\theta$  parameters computed from similarities and the quantum  $\theta$  parameters corresponding to the interference terms in the Quantum-Like Bayesian Network model. This assumption comes from the book of Busemeyer and Bruza (2012), where it is stated that the  $\theta$  parameter that arises in quantum interference effects corresponds to the phase of the angle of the inner product between the projectors of two random variables. They also state that the inner product provides a measure of similarity between two vectors (where each vector corresponds to a superposition of events). If the vectors are unit length, then the Cosine Similarity collapses to the inner product. Given all these relations, we can assume that the similarities computed between two vectors representing the images of faces can be used to set quantum interference parameters.

## Acknowledgments

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013 and through the Ph.D. grant SFRH/BD/92391/2013. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We would also like to thank Dr. Jerome Busemeyer and Dr. Zheng Joyce Wang for providing us the dataset of images used in their experiments.

## References

- Asano, M., Basieva, I., Khrennikov, A., Ohya, M., Tanaka, Y., & Yamato, I. (2012). Quantum-like model for the adaptive dynamics of the genetic regulation of *E. coli*'s metabolism of glucose/lactose. *Journal of Systems and Synthetic Biology*, 6, 1–7.
- Asano, M., Khrennikov, A., & Ohya, M. (2015). *Quantum adaptive biology: from genetics to cognition*. Springer.
- Brandenburger, A. (2010). The relationship between quantum and classical correlation in games. *Games and Economic Behavior*, 69, 175–183.
- Busemeyer, J., & Bruza, P. (2012). *Quantum model of cognition and decision*. Cambridge University Press.
- Busemeyer, J., Wang, Z., & Lambert-Mogiliansky, A. (2009). Empirical comparison of Markov and quantum models of decision making. *Journal of Mathematical Psychology*, 53, 423–433.
- Busemeyer, J., Wang, Z., & Shiffrin, R. (2015). Bayesian model comparison favors quantum over standard decision theory account of dynamic inconsistencies. *Decision*, 2, 1–12.
- Conte, E. (2008). Testing quantum consciousness. *NeuroQuantology*, 6, 126–139.
- Conte, E., Todarello, O., Federici, A., Vitiello, F., Lopane, M., Khrennikov, A., & Zbilut, J. (2007). Some remarks on an experiment suggesting quantum like behavior of cognitive entities and formulation of an abstract quantum mechanical formalism to describe cognitive entity and its dynamics. *Chaos, Solitons & Fractals*, 31, 1076–1088.
- Deutsch, D. (1988). Quantum theory of probability and decisions. *Proceedings of the Royal Society of London, Series A (Mathematical and Physical Sciences)*, 455, 3129–3137.
- Haven, E., & Khrennikov, A. (2013). *Quantum social science*. Cambridge University Press.
- Kahn, C. E., Roberts, L. M., Shaffer, K. A., & Haddawy, P. (1997). Construction of a Bayesian network for mammographic diagnosis of breast cancer. *Computers in Biology and Medicine*, 27, 19–29.
- Khrennikov, A. (2009). Classical and quantum-like randomness and the financial market. In *Coping with the complexity of economics*. Springer.

- 1 Khrennikov, A., & Basieva, I. (2014). Possibility to agree on disagree from quantum  
information and decision making. *Journal of Mathematical Psychology*, 62–63,  
1–15.
- 2 Khrennikov, A., Basieva, I., Dzharfarov, E., & Busemeyer, J. (2014). Quantum models  
for psychological measurements: An unsolved problem. *PLoS One*, 9, 1–8.
- 3 Khrennikov, A., & 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, 378–388.
- 4 Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and  
techniques*. The MIT Press.
- 5 Martínez-Martínez, I., & Sánchez-Burillo, E. (2016). Quantum stochastic walks on  
networks for decision-making. *Scientific Reports*.
- 6 Moreira, C., & Wichert, A. (2014). Interference effects in quantum belief networks.  
*Applied Soft Computing*, 25, 64–85.
- 7 Moreira, C., & Wichert, A. (2015a). The relation between acausality and interference  
in quantum-like Bayesian networks. In *Proceedings of the 9th international  
conference on quantum interactions*.
- 8 Moreira, C., & Wichert, A. (2015b). The synchronicity principle under quantum  
probabilistic inferences. *NeuroQuantology*, 13, 111–133.
- 9 Moreira, C., & Wichert, A. (2016). Quantum-like Bayesian networks for modeling  
decision making. *Frontiers in Psychology*, 7.
- 10 Mura, P. L. (2005). Correlated equilibria of classical strategic games with quantum  
signals. *International Journal of Quantum Information*, 3, 183–188.
- 11 Murphy, K. (2012). *Machine learning: a probabilistic perspective*. MIT Press.
- 12 Pothos, E., & Busemeyer, J. (2009). A quantum probability explanation for violations  
of rational decision theory. *Proceedings of the Royal Society B*, 276, 2171–2178.
- 13 Pothos, E., & Busemeyer, J. (2013). Can quantum probability provide a new direction  
for cognitive modeling? *Brain and Behavioral Science*, 36, 255–327.
- 14 Pothos, E., Busemeyer, J., & Trueblood, J. (2013). A quantum geometric model of  
similarity. *Psychological Review*, 120, 679–696.
- Russel, S., & Norvig, P. (2010). *Artificial intelligence: a modern approach* (3rd ed.).  
Pearson Education.
- Sahami, M., Dumais, S., Heckerman, D., & Horvitz, E. (1998). A Bayesian approach to  
filtering junk e-mail. In *Proceedings of the 15th national conference on artificial  
intelligence, workshop on learning for text categorization*.
- Townsend, J., Silva, K., 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.
- Trueblood, J., & Busemeyer, J. (2011). A comparison of the belief-adjustment model  
and the quantum inference model as explanations of order effects in human  
inference. *Cognitive Science*.
- Trueblood, J., Pothos, E., & Busemeyer, J. (2014). Quantum probability theory as a  
common framework for reasoning and similarity. *Frontiers in Psychology*, 5.
- Tucci, R. (1995). Quantum Bayesian nets. *International Journal of Modern Physics B*,  
9, 295–337.
- van Rijsbergen, C. J. (2004). *The geometry of information retrieval*. Cambridge  
University Press.
- Wang, Z., & Busemeyer, J. (2016). Interference effects of categorization on decision  
making. *Cognition*, 150, 133–149.
- Wang, Z., Solloway, T., Shiffrin, R., & Busemeyer, J. (2014). Context effects produced  
by question orders reveal quantum nature of human judgments. *Proceedings of  
the National Academy of Sciences*, 111, 9431–9436.
- Yearsley, J., & Pothos, E. (2014). Challenging the classical notion of time in cognition:  
a quantum perspective. *Proceedings of the Royal Society B*.
- Yukalov, V., & Sornette, D. (2010). Entanglement production in quantum decision  
making. *Physics of Atomic Nuclei*, 73, 559–562.
- Yukalov, V., & Sornette, D. (2011). Decision theory with prospect interference and  
entanglement. *Theory and Decision*, 70, 283–328.
- Zou, M., & Conzen, S. (2005). A new dynamic Bayesian network (dbn) approach  
for identifying gene regulatory networks from time course microarray data.  
*Bioinformatics*, 21, 71–79.