

Applying quantum principles to psychology

Jerome R Busemeyer¹, Zheng Wang², Andrei Khrennikov³ and Irina Basieva³

¹ Department of Psychology and Brain Science, Indiana University, Bloomington, USA

² Center for Cognitive and Brain Science, The Ohio State University, Columbus, OH 43210, USA

³ International Center for Mathematical Modeling in Physics, Engineering, Economics, and Cognitive Science, Linnaeus University, Växjö-Kalmar, S-35195, Sweden

E-mail: Andrei.Khrennikov@lnu.se

Received 23 May 2014, revised 11 August 2014

Accepted for publication 29 September 2014

Published 19 December 2014

Abstract

This article starts out with a detailed example illustrating the utility of applying quantum probability to psychology. Then it describes several alternative mathematical methods for mapping fundamental quantum concepts (such as state preparation, measurement, state evolution) to fundamental psychological concepts (such as stimulus, response, information processing). For state preparation, we consider both pure states and densities with mixtures. For measurement, we consider projective measurements and positive operator valued measurements. The advantages and disadvantages of each method with respect to applications in psychology are discussed.

Keywords: quantum probability, quantum(-like) cognition and psychology, order effects, violation of the law of total probability, quantum measurement approach to cognition, projective measurements, positive operator valued measures

1. Why apply quantum theory to psychology?

Twenty years ago, a group of physicists and psychologists introduced the bold idea of applying the abstract principles from quantum theory outside of physics to the field of human judgment and decision making [1–4]. In contrast to some recent proposals concerning the quantum physics of the brain, e.g., [5–7], this new framework does *not* rely on the assumption that the brain is some type of quantum computer, and instead it uses a probabilistic formulation borrowed from quantum theory that involves non-commutative algebraic principles [8–11]. This new field, called *quantum cognition*, has proved to be able to account for puzzling behavioral phenomena that are found in studies of a variety of human judgments and decisions including violations of the ‘rational’ principles of decision making [12], conjunction and disjunction probability judgment errors [13], over- and under-extension errors in conceptual combinations [14, 15], ambiguous concepts [16], order effects on probabilistic inference [17, 18], interference of categorization on decision making [19], attitude question order effects [20] and other puzzling results from decision research [21–24]. In short, quantum models of judgment and decision have made

impressive progress organizing and accounting for a wide range of puzzling findings using a common set of principles.

1.1. Example: Categorization-decision (C-D) experiment

To see more concretely how quantum theory can be applied to psychology, consider the following psychology experiment used to investigate the interference of categorization on decision making. Often decision makers need to make categorizations before choosing an action. For example, a military operator has to categorize an agent as an enemy before attacking with a drone. How does this overt report of the category affect the later decision? This paradigm was originally designed to test a Markov model of decision making that is popular in psychology [25]. Later it was adapted to investigate ‘quantum like’ interference effects in psychology [19].

We begin by briefly summarizing the methods used in the experiments. On each trial of several hundred training trials, the participant is first shown a picture of a face that may belong to a ‘good guy’ category (category G) or a ‘bad guy’ category (category B), and they have to decide whether to ‘attack’ (action A) or ‘withdraw’ (action W). The trial ends with feedback indicating the category and appropriate action that was assigned to the face on that trial. There are many

different faces, and each face is probabilistically assigned to a category, and the appropriate action is probabilistically dependent on the category assignment. Some of the faces are usually assigned to the ‘good guy’ category, while other faces are usually assigned to the ‘bad guy’ category. The category is important because participants are usually rewarded (win points worth money) for ‘attacking’ faces assigned to ‘bad guys’ and they are usually punished (lose points worth money) for ‘attacking’ faces assigned to the ‘good guys;’ likewise they are usually rewarded for ‘withdrawing’ from ‘good guys’ and punished for ‘withdrawing’ from ‘bad guys.’ Participants are given ample training during which they learn to first categorize a face and then decide an action, and feedback is provided on both the category and the decision. Although the feedback given at the end of each trial is probabilistic, the optimal decision is to always ‘attack’ when the face is usually assigned to a ‘bad guy’ category, and always ‘withdraw’ when the face is usually assigned to a ‘good guy’ category. The key manipulation occurs during a transfer test phase which includes the standard C-D trials followed by either ‘category alone’ (C-alone) trials or ‘decision alone’ (D-alone) trials. For example, on a ‘decision alone’ trial, the person is shown a face, and simply decides to ‘attack’ or ‘withdraw,’ and receives feedback on the decision. The categorization of the face on the D-alone trials remains just as important to the decision as it is on C-D trials, and some implicit inference about the category is necessary before making the decision, but the person does not overtly report this implicit inference.

Note that the C-D condition in the psychology experiment allows the experimenter to observe which ‘path’ the participant follows before reaching a final decision. This is analogous to a ‘double slit’ physics experiment in which the experimenter observes which ‘path’ a particle follows before reaching a final detector. In contrast, for the D-alone condition in the psychology experiment, the experimenter does not observe which ‘path’ the decision maker follows before reaching a final decision. This is analogous to the ‘double slit’ physics experiment in which the experimenter does not observe which ‘path’ the particle follows before reaching a final detector⁴.

According to the Markov model proposed in [25], for the D-alone condition, the person implicitly performs the same task as explicitly required by the C-D condition. More specifically, for the D-alone condition, once a face (denoted f) is presented, there is a probability that the person implicitly categorizes the face as a ‘good’ or ‘bad’ guy. From each category inference state, there is a probability of transiting to the ‘attack’ or ‘withdraw’ decision state. So the probability of ‘attack’ in the D-alone condition (denoted as $p(A|f)$) should equal the total probability of ‘attacking’ in the C-D condition (denoted as $p_T(A|f)$). The latter is defined by the probability that the person categorizes a face as a ‘good guy’ and then ‘attacks’ plus the probability that the

person categorizes the face as a ‘bad guy’ and then ‘attacks’ ($p_T(A|f) = p(G \cap A|f) + p(B \cap A|f)$). Using this C-D paradigm, one can examine how the overt report of the category interferes with the subsequent decision. An *interference effect* of categorization on decision making occurs when the probability of ‘attacking’ for D-alone trials differs from the total probability pooled across C-D trials. The Markov model for this task originally investigated by [25] predicts that there should be no interference, and the law of total probability should be satisfied.

Beginning with our first study [19], we have conducted a series of four experiments on this paradigm. The results of these experiments all generally show the same results, but we briefly report a summary of findings from the fourth experiment that included 246 participants (a minimum 34 observations per person per condition). When a face most likely is assigned to the ‘good guy’ category (we denote these faces as g), the law of total probability is approximately satisfied ($p_T(A|g) = .36$, $p(A|g) = .37$). However, when a face most likely is assigned to the ‘bad guy’ category (we denote these faces as b), the probability of ‘attack’ (i.e. the optimal decision with respect to the average payoff) is systematically greater for the D-alone condition as compared to the C-D condition’ violating the law of total probability ($p(A|b) = .62 > p_T(A|b) = .56$)⁵. More surprising, the probability of ‘attack’ for the D-alone condition (which leaves the ‘good’ or ‘bad’ guy category unresolved) was even greater than the probability of ‘attack’ given that the person previously categorized the face as a ‘bad guy’ ($p(A|b) = .62 > p(A|b, B) = .61$) on a C-D trial! For some reason, the overt categorization response interfered with the decision by reducing the tendency to ‘attack’ faces that most likely belonged to the ‘bad guy’ category. These violations of the law of total probability run counter to the predictions of the Markov model proposed by [25] for this task.

1.2. A Quantum decision model

The details of a quantum model for the C-D task are presented in [19], and here we only present a brief summary. The human decision system is represented by a unit length state vector $|\psi\rangle$ that lies within an four-dimensional Hilbert space spanned by four basis vectors. Each basis vector represents one of the four combinations of categories and actions (e.g., $|GA\rangle$ is a basis vector corresponding to category G and action A). The state $|\psi_f\rangle = \psi_{GA} |GA\rangle + \psi_{GW} |GW\rangle + \psi_{BA} |BA\rangle + \psi_{BW} |BW\rangle$ is prepared by the face stimulus f that is presented on a trial. The question about the category is represented by a pair of projectors for good and bad categories $C_G = |GA\rangle\langle GA| + |GW\rangle\langle GW|$, $C_B = (I - C_G)$. The question about the action is represented by a pair of projectors for attack and withdraw actions $D_A = U_{DC} |GA\rangle\langle GA| U_{DC}^\dagger + U_{DC} |BA\rangle\langle BA| U_{DC}^\dagger$, $D_W = (I - D_A)$, where U_{DC} is a unitary operator that changes the basis from the categorization to the decision basis.

⁴ We remark that here the picture of path is used only for illustrative purpose; therefore we placed path in quotation marks. In QM there is no such concept as a ‘path’ (trajectory) of a particle. We can only ascertain, and then only statistically, a singular event of an electron ‘passing’ through a slit. In fact this way of seeing the situation even provided a better parallel here.

⁵ This difference are statistically significant: $t(245) = 4.41$, $p = .0004$. Also this same effect was replicated in four independent experiments.

The probability of first categorizing the face as a ‘bad guy’ and then ‘attacking’ equals $p(B, A|f) = p(B) \cdot p(A|B) = \|C_B|\psi_f\rangle\|^2 \cdot \|D_A|\psi_B\rangle\|^2$, with $|\psi_B\rangle = \frac{C_B|\psi_f\rangle}{\|C_B|\psi_f\rangle\|}$, and combining the terms in the product we obtain $p(B, A|f) = \|D_A \cdot C_B \cdot |\psi_f\rangle\|^2$; similarly, the probability of first categorizing the face as a ‘good guy’ and then ‘attacking’ equals $p(G, A|f) = \|D_A \cdot C_G \cdot |\psi_f\rangle\|^2$; and so the total probability of attacking under the C-D condition equals $p_T(A|f) = \|D_A \cdot C_G \cdot |\psi_f\rangle\|^2 + \|D_A \cdot C_B \cdot |\psi_f\rangle\|^2$. The probability of attack in the D-alone condition equals $p(A|f) = \|D_A \cdot |\psi_f\rangle\|^2 = \|D_A \cdot (C_G + C_B)|\psi_f\rangle\|^2 = \|D_A \cdot C_G|\psi_f\rangle + D_A \cdot C_B|\psi_f\rangle\|^2 = \|D_A \cdot C_G|\psi_f\rangle\|^2 + \|D_A \cdot C_B|\psi_f\rangle\|^2 + \text{Int}$, where $\text{Int} = 2 \cdot \text{Re}[\langle \psi_f | C_G D_A C_B | \psi_f \rangle]$. If the projectors for categorization commute with the projectors for action (e.g., $U_{DC} = I$), then the interference is zero, $\text{Int} = 0$, and we obtain $p(A|f) = \|D_A \cdot C_G|\psi_f\rangle\|^2 + \|D_A \cdot C_B|\psi_f\rangle\|^2 = p_T(A|f)$, and the law of total probability is satisfied. However, if the projectors do not commute (e.g., $U_{DC} \neq I$), then we obtain an interference term. We can select the unitary operator U_{DC} to produce an inner product $\text{Int} = -.06$, and account for the observed violation of the law of total probability.

We originally conducted these experiments because we predicted that an interference effect of categorization on decisions would occur based on past research using quantum models of decision [26]. However, we could not predict the direction or quantitative size of the interference. Now that we have this estimate, we can use it to make new predictions for new experiments. Along this line, we carried out a second condition in our fourth study to test our model. During the transfer phase of the second condition, we included two different types of transfer test trials: (a) D-alone trials as described earlier in which the participant did not categorize but only made an action decision, and (b) trials on which we informed the participant about the correct category and the person made an action decision. Note that for the second condition (b), the participants did not make any categorization response and only made a decision, but this decision was based on information provided by the experimenter about the category. According to our theory, providing information about the category produces the same effect as taking a measurement of the category—if the person is told the face belongs to the ‘bad guy category,’ then the state is updated from $|\psi_f\rangle$ to $|\psi_B\rangle = \frac{C_B|\psi_f\rangle}{\|C_B|\psi_f\rangle\|}$ in the same way as if a measurement was made. Therefore, if we use the choice probabilities from this second condition to compute the total probability, then this second condition should produce exactly the same interference effect as the first. Supporting this prediction, the interference term for the second condition equaled $\text{Int} = -.05$, which closely approximates the interference obtained from the first condition. The great challenge for quantum cognition models is to use the same principles then to predict new findings (e.g., see [20, 27], further *a priori* tests of the theory).

2. Step by step application of quantum theory to psychology

How does one apply quantum principles to psychological experiments? How are the stimuli and responses of a psychology experiment related to the state preparation and measurements of quantum theory? How does one determine the observables, initial states, and unitary operators for psychology? The above application involved several important assumptions regarding the mapping of basic concepts in physics (e.g., state preparation, state evolution, measurement operators) into basic concepts of psychology (e.g., stimuli, information processing, responses). The next sections examine these mappings more carefully, explore different ways to formulate this mapping, and discuss the advantages and disadvantages of different mappings.

2.1. Choosing a Hilbert space

Both physicists and psychologists are faced with the scientific task of making predictions about the probability that different kinds of events occur in their experiments. For example the psychologist wants to predict whether a person will ‘attack’ or ‘withdraw’ when a face is presented; a quantum physicist wants to predict whether a particle is detected by one detector or another after it is emitted from a source. Traditionally, psychologists have used classic theory (axiomatized by Kolmogorov), whereas quantum physicists use quantum theory (axiomatized by Dirac and von Neumann) to make these predictions.

Classic theory represents events as subsets of a universal set called the sample space. Quantum theory represents events as subspaces of universal vector space called the Hilbert space. Technically, a Hilbert space is a vector space defined on a complex field and endowed with an inner product that satisfies completeness. The adoption of a vector space instead of a sample space to represent events may be the most important assumption that is made in the application of quantum theory to psychology.

Physicists work with both finite and infinite dimensional Hilbert spaces. For example, an infinite dimensional space is used to represent the position of a particle, but a finite n -dimensional space is used to represent the spin of a particle. There is no *a priori* reason why a psychologist could not use infinite dimensional spaces, but as in the field of quantum computing, almost all of the previous work in quantum cognition has used finite dimensional spaces. Choosing the dimensionality of the Hilbert space is critical first step for constructing a quantum cognition model. The dimension depends on the number of variables and the number of values on each variable.

Two issues need to be considered when choosing the dimension of the Hilbert space. The first issue is whether the measurements of a variable are considered coarse (degenerate) or complete (non-degenerate) (see, e.g., [28]). The second issue arises when there are two or more variables and the question is whether or not the variables should be combined to form what is called a tensor product space.

First consider the issue of completeness. The measurement of a variable is complete if the outcomes cannot be refined, and so each outcome can be represented by a single ray or a single dimension. Consider an example such as asking a juror to rate the probability of guilt on a five level scale (1 = very low probability, 2 = moderately low probability, 3 = uncertain, 4 = moderately high probability, 5 = very high probability). This variable is measured by five mutually exclusive and exhaustive outcomes and so one might wish to represent this variable by a five dimensional space. However, a person may be capable of rating confidence on a much finer scale. Suppose the finest scale is a 21 level scale (0 = certainly not guilty, 10, ..., 50 = equally likely, 60, ..., 90, 100 = certain guilty). Then the 21 level scale forms a complete measurement and requires a 21 dimensional space. If the experimenter uses a five level scale, then this represents a coarse measurement defined within a higher 21 dimensional space (e.g., levels 70–100 could be mapped into category 5). In reality, we do not know how fine is the internal scale of the human decision maker.

Next consider the tensor product space issue. Often in psychology experiments, the participant is asked more than one question. Consider our previous example of the C-D experiment. In this study, there are two variables: one is the categorization (categorize as good versus bad) and the other is the action (choose between attack versus withdraw). The problem is how to represent both variables within a single vector space. The simplest possible representation is obtained by using the same two dimensional space to represent both variables. This is done by using a different basis to represent each variable within the same two dimensional space. A more complex representation is obtained by using a four dimensional space that represents each combination of values for the two variables. The latter is called a tensor product space.

How does one decide whether to represent the two variables within the same space or to combine the variables into a tensor product state? The tensor product representation assumes that it is possible for a person to *simultaneously* consider the values of *both* variables at the same time. The consideration of one value on one variable does not disturb the evaluation of the second variable, and there should be no order effects when the two variables are considered *simultaneously*⁶. If this simultaneous evaluation is not possible, and the consideration of the value of one variable disturbs the evaluation of the other variable, producing order effects, then the variables have to be evaluated sequentially by changing the basis within a common space.

Ultimately, picking the dimension becomes an empirical question. This must be done by balancing accuracy and parsimony. One starts with the simplest model, and if that fails empirically, then one is forced to gradually increase complexity. In [19] we initially tried the simplest two dimensional space with different bases for each variable, but this failed empirically. Then we tried the next simplest representation

based a four dimensional (tensor product) vector space, and this was empirically satisfactory.

2.2. Choosing a basis to construct projectors and observables

A basis is a set of orthonormal basis vectors that span the Hilbert space. Once we select a basis, we can construct a projector for the subspace representing an event by using its basis vectors. The projector for a subspace is formed by the sum of the outer products of basis vectors that span the subspace. Consider once again the C-D task, and suppose that the four orthonormal vectors ($|GA\rangle, |GW\rangle, |BA\rangle, |BW\rangle$) form the basis for the categorization. These four vectors can be represented numerically by $[1, 0, 0, 0]'$, $[0, 1, 0, 0]'$, $[0, 0, 1, 0]'$, $[0, 0, 0, 1]'$, respectively. If we wish to measure the event 'bad guy category' then we choose the basis vectors $|BA\rangle, |BW\rangle$ representing this category, and form the projector $C_B = |BA\rangle\langle BA| + |BW\rangle\langle BW|$, which numerically corresponds to the matrix $\text{diag}[0, 0, 1, 1]$. Likewise we can define the projector for the 'good guy' category using the basis vectors $|GA\rangle, |GW\rangle$, and numerically this projector corresponds to the matrix $\text{diag}[1, 1, 0, 0]$. If a set of projectors are pairwise orthogonal and sum to the identity, then it forms a complete set that represents a mutually exclusive and exhaustive set of events. In this example, the two projectors are orthogonal, $C_G C_B = 0$, and they sum to the identity, $C_G + C_B = I$, and so they form a complete set.

Using the projectors of a complete set, we can form an observable by assigning a real number to each projector. For example, suppose we assign +1 to the 'bad guy' event and -1 to the 'good guy' event; then the observable for the categorization variable is defined as $C = (-1) \cdot C_G + (+1) \cdot C_B = C_B - C_G$. This observable is constructed from the four categorization ($|GA\rangle, |GW\rangle, |BA\rangle, |BW\rangle$) basis vectors. These four basis vectors are eigenvectors of this observable because each one satisfies the eigenvector equation, for example $C|BW\rangle = (+1) \cdot |BW\rangle$. The same eigenvalue (+1) is assigned to both eigenvectors $|BA\rangle, |BW\rangle$. A repeated eigenvalue like this is called a degenerate eigenvalue—any linear combination of eigenvectors $|BA\rangle, |BW\rangle$, which share the same (+1) eigenvalue is also an eigenvector of C with the same (+1) eigenvalue. A degenerate eigenvalue implies that the observation of that value is a coarse measurement because there is more than one basis vector associated with that value.

There is, however, an infinite number of choices for the basis. Different questions might require a different choice of basis to represent the answers. Consider once again the categorization decision task described earlier. One basis could be used to judge the strength of evidence favoring each category (good versus bad categories), but a different basis may be needed to evaluate the consequences of actions (attack versus withdraw actions). In the above example, the unitary operator U_{DC} was used to change the basis from the one used for categorization to the one used for action. The new basis for the action decision is ($U_{DC}|GA\rangle, U_{DC}|GW\rangle, U_{DC}|BA\rangle, U_{DC}|BW\rangle$). Then the projector for the 'attack action' is $D_A =$

⁶ Simultaneity is an important condition, because if a unitary transformation occurs in between the measurements, then order effects can occur.

$U_{DC}(|GA\rangle\langle GA| + |BA\rangle\langle BA|)U_{DC}^\dagger$. Likewise the projector for the ‘withdraw’ action is $D_W = (I - D_A)$. If we assign -1 to the ‘withdraw’ action and $+1$ to the ‘attack’ action, then the observable is $D = D_A - D_W$. Note that the observable for the action does not commute with the observable for the categorization because the commutator $DC - CD \neq 0$ does not equal zero.

The most difficult task is determining the form of the unitary operator, U_{DC} , used to change the basis. A completely general form for a unitary operator to a basis $Y = |Y_i\rangle, i = 1, N$ from another basis X is $U_{YX} = \sum_j |Y_j\rangle\langle X_j|$, but this does not give much guidance. One general way to construct a unitary matrix is to use the exponential function of the operator H , $U_{DC} = \exp(-iH)$, where H is a self adjoint linear operator ($H = H^\dagger$) called the Hamiltonian. Any unitary operator can be formed in this manner. Using this method, the problem then becomes one of choosing the form of the Hamiltonian. In [19], we designed a specific Hamiltonian based on the rewards and punishments of the categorization–decision task.

If we present different faces but ask the same questions about categories and actions for each face, then the same observables C, D can be applied to each face. Thus, according to this interpretation, the observable only depends on the question and it does not depend on the stimulus. We should point out, however, that the separation between state preparation and selection of the observable is not always clear cut, and one could argue that they can’t be separated, so there are alternative viewpoints on this issue that we discuss later.

2.3. Preparing states

In quantum theory, the system under investigation is represented as a unit length vector, $|\psi\rangle$, in the Hilbert space. In physics the system often refers to a particle, but in psychology the system usually refers to a person. The state vector $|\psi\rangle$ can be expressed as a linear combination of the basis vectors and the coordinates of the state vector with respect to a basis are called the amplitudes. For example, referring back to the C-D task, if we choose the categorization basis, then the state vector is defined by $|\psi\rangle = \psi_{GA}|GA\rangle + \psi_{GW}|GW\rangle + \psi_{BA}|BA\rangle + \psi_{BW}|BW\rangle$, and ψ_{GW} is for example the amplitude assigned to $|GW\rangle$. Generally, these amplitudes can be complex numbers and the sum of the squared magnitudes equals one. The interpretation of these complex numbers is difficult for many psychologists, but there is no *a priori* reason for limiting psychological applications to real numbers, just as there is no reason to limit electrical engineering applications to real numbers. In fact, Fourier analysis, using the complex transform, is commonly used in both electrical control engineering, neural signal processing, and human psychophysics. Ultimately the answers that we obtain and need to interpret are always real.

In physics, the experiment begins with some physical system in some state and then the experimenter prepares the state of the system by applying physical devices before testing begins. Different types of physical tests can be

performed on systems after they are prepared in the same state. In psychology, the person begins with some state, and then the experimenter manipulates the state of the person by presenting information or a stimulus prior to questioning. In our C-D experiment, the participant is presented a new face on each trial, and the experimenter asks questions about the category and the action to take for that face. Different types of questions (which category, which action) can be asked about the same face. Therefore the state before questioning is conditioned on the stimulus, $|\psi_f\rangle$, where f indicates that the person was shown a face labeled f before asking any questions. Thus, according to this interpretation, the presentation of a stimulus in a psychology experiment corresponds to state preparation in a physics experiment and the state (before the question) does not depend on the question that is asked later. Any question can be asked from this state, and the probability of answers will vary across questions for the same state. Once again we should point out that this separation between state preparation and selection of the observable is debatable, and later we consider alternative viewpoints on this issue.

There are two different ways to prepare the state in quantum physics. One is by a measurement that projects the state to a subspace followed by normalization to unit length. The other is by application of a unitary operator that ‘rotates’ the state in the Hilbert space while maintaining unit length. Both of these methods could be used in a psychology experiment.

There is also a thorny problem of interpretation here, as concern the status of ‘quantum state’ as a mathematical object, a state vector that is, whether it corresponds to any actual reality, say, a physical state of a particle in quantum physics, or only provides, as Schrödinger called, a probability catalog for future predictions concerning possible experiments. During the last hundred years of stormy debates on quantum foundations this problem, the problem of the interpretation of the wave function, was not solved. The quantum(-like) cognition approach borrowed not only advances of QM, but also all its problems and, in particular, this problem. Is it really the belief state of an individual person? Is it just a probability catalog? For a moment, we cannot provide the definite answer and we proceed pragmatically by using the mathematical formalism of QM for analysis of statistical data, i.e., in fact, we interpret the belief state as a probability catalog. However, we cannot exclude that the ‘individual interpretation’ of the belief state can also be justified.

The stimulus information that the participant experiences changes the state of the person. For example, in the C-D experiment, the face that is presented at the beginning of a trial will influence the state of the person before making a categorization. For example, if the face looks like a ‘bad guy’ the state will move toward the subspace for that category. There are various ways that this change could happen. One way is to use a unitary operator U_f that depends on the face stimulus. If the initial state, before the face is presented, is

defined as $|\psi_I\rangle$, then the state becomes $|\psi_f\rangle = U_f |\psi_I\rangle$ after presentation of the face⁷.

The experimenter could present the participant some facts, which if accepted to be true, would cause the person's state to be projected onto the subspace consistent with those facts, and normalized to have unit length. For example, in the C-D experiment, the participant could be informed that the face actually belongs to the 'bad guy' category before being asked to make an action decision. Define $|\psi_f\rangle$ as the state after seeing the face but before any category information is presented on a trial. After the new category information is presented, the state is updated to become $|\psi_B\rangle = (1/c) \cdot C_B |\psi_f\rangle$, $c = \|C_B |\psi_f\rangle\|$ and then the decision to attack is based on this updated state.

2.4. Computing probabilities and updating states

The purpose of using quantum theory for both physicists and psychologists is to predict the probability of events. For a given state, the probability of an event is obtained by projecting the state vector onto the subspace for the event and computing its squared length. For example, referring again to the C-D study, if the person sees a face and is asked to categorize it, then the probability of 'bad guy' equals $p(B|f) = \|C_B |\psi_f\rangle\|^2$. In physics, the state of the system changes following a measurement. The same process occurs in psychology—asking a question and deciding on a definite answer changes the state of the person. According to Lüder's rule, if the person categorizes the face as a 'bad guy,' then the new state equals $|\psi_B\rangle = (1/c) \cdot C_B |\psi_f\rangle$, $c = \|C_B |\psi_f\rangle\|$. From this it follows that the probability of categorizing as a 'bad guy' and then deciding to 'attack' equals $p(B, A|f) = \|D_A \cdot |\psi_B\rangle\|^2 \cdot \|C_B \cdot |\psi_f\rangle\|^2 = \|D_A \cdot C_B \cdot |\psi_f\rangle\|^2$. At the very end of the trial, with acategorization of 'bad guy' and the decision to 'attack' for example, the final state becomes $|\psi_{AB}\rangle = (1/d) \cdot D_A |\psi_B\rangle$, where $d = \|D_A |\psi_B\rangle\|$.

2.5. Sequential effects

An important question that still needs to be addressed concerns the changes in the state of the person from one trial to the next. For example, in the C-D paradigm, each trial begins with the presentation of a face and ends with feedback about the correct category and action. This trial structure is part of the instructions given to the participant at the beginning of the experiment. In this way, the person is trained with feedback on the probabilities of faces being assigned to categories and appropriate actions. How can this change from learning by feedback be incorporated into the system? There are many ways to do this, but one way is to use the feedback to update the Hamiltonians that are used to form the unitary transformations. For example, in the category-decision making task, the unitary matrix U_f in the transformation $|\psi_f\rangle = U_f |\psi_I\rangle$ can

be updated through feedback about the category. Also the unitary matrix U_{DC} that changes the basis from categorization to decision can be updated based on feedback about the correct action.

A separate question concerns possible carry over effects from answers on one trial to the next trial. In other words, how does the final state at the end of one trial, say $|\psi_{AB}, t\rangle$ after categorizing a face as a 'bad guy' and deciding to 'attack', evolve during the inter-trial interval into the initial state $|\psi_I, t + 1\rangle$ before the face is presented for the next trial? Usually the information accumulated on one trial is not relevant for the next trial, and the participants are instructed to treat the trials separately and independently. The inter-trial interval separating trials is made clear to the participant and sufficiently long to prepare for the next trial (e.g. a reasonable pause with a blank screen). The state needs to be reset during this inter-trial interval from some final state (e.g., $|\psi_{AB}, t\rangle$) after the previous trial back to a common neutral state $|\psi_I, t + 1\rangle = |\psi_0\rangle$ before the next trial begins. To accomplish this reset task, some unitary operation or projection is required to change the state during the inter-trial interval. This process is an important issue but little is understood about it. By the way, the same issue arises with more traditional cognitive models such as Markov models, and so the problem is not unique to quantum models.

In some experiments, however, the information accumulated during one episode remains relevant for another episode and so both episodes together form a trial. In this case, there is little or no inter-trial interval separating the episodes. If a person is shown a face and asked to categorize it with respect to aggressiveness, and the same face is continued to be shown but now the person is asked to categorize it with respect to intelligence, and the trial is defined by the pair of episodes, then the participant can connect these two episodes together so that the state following the answer to the first question (aggressiveness) is carried over and used as the state for answering the second question (intelligence). In other words, no unitary transform to reset the system intervenes between the two questions. The experimental instructions and conditions that determine whether the reset versus the carry over occurs is an important matter for future research [29].

2.6. Positive operator value measures

So far we have limited our discussion to measurement defined as projectors which satisfy $D_A = D_A^\dagger = D_A^2$. A more general type of measurement is one that does not need to satisfy either of these two properties, and instead only satisfies positivity and completeness properties, defined below. These generalized measurements form what are called the positive operator value measurements (POVM's). As recommended in [29], POVM's could be an important tool for psychologists who need to work with more complex types of measurements. As an example, consider the linear operator $P_{AB} = D_A C_B$, which is a measurement of the sequence of events 'categorize as bad guy' and then 'attack.' If the projectors D_A, C_B commute, then P_{AB} defines the projector for the conjunction of these two events, which is represented as the intersection of subspaces

⁷ To be more precise, we should use the notation $|\psi_{I,f}\rangle = U_f |\psi_I\rangle$, because the transformed state also depends on the initial state. But to avoid using too many subscripts, and keeping in mind the history, hereafter we use the shorter notation.

A, B . Of course, this implies no order effects. If there are order effects, then the events D_A, C_B do not commute and the conjunction (which is commutative) is not defined. For the non-commutative case, the linear operator P_{AB} is not a projector, it does not correspond to any single subspace, and it is not an event. Instead this measurement operator represents a sequence of two events. There are three other measurement operators for the other three sequences of events in this task: $P_{AG} = D_A C_G, P_{WB} = D_W C_B, P_{WG} = D_W C_G$. Note that $P_{AB}^\dagger P_{AB}$ is a positive operator because $\langle \psi | P_{AB}^\dagger P_{AB} | \psi \rangle \geq 0$ for all $|\psi\rangle$, and the sum $P_{AB}^\dagger P_{AB} + P_{AG}^\dagger P_{AG} + P_{WB}^\dagger P_{WB} + P_{WG}^\dagger P_{WG} = I$ satisfies completeness. (Technically, P_{AB} is the measurement operator and $P_{AB}^\dagger P_{AB}$ is the positive operator corresponding to this measurement). These two properties guarantee that the probabilities of these four mutually exclusive and exhaustive sequences to sum to unity. Positivity and completeness are the two requirements needed to define a complete set of POVM's. The four measurement operators provide one way to model sequences of events when there is order dependence. A more general way to model conjunctions of events in psychology using a POVM formulation was recently proposed by [30].

Applications of POVMs to the problems of decision making in the framework of theory of open quantum systems were considered in [31]. General discussion on a possibility to describe cognitive phenomena solely by using only observables represented by Hermitian operators can found in [32]. Here it was shown that some statistical data from cognitive psychology cannot be represented with the aid of Hermitian operators; one has to use POVMs and even their generalizations.

2.7. Constructing mixed states

The state vector $|\psi\rangle$ is called a pure state. This pure state can also be expressed as a density operator $\rho = |\psi\rangle\langle\psi|$ and then the probability of an event, such as to categorize a face as 'bad guy,' is given by the rule $p(C_B) = \text{tr}[C_B \cdot \rho]$, which is equivalent to the previously defined rule. However, this representation of state allows one to define a more general mixture state $\rho = \sum p_j |\psi_j\rangle\langle\psi_j|, 0 \leq p_j \leq 1, \sum p_j = 1$, and apply the same rule for computing probabilities from densities. The advantage of using the density operator is that this more general form can be used to represent a probability obtained from a mixture of participants, where each participant is represented by a different pure state [33]. One complicating factor that arises when working mixed states is that their decomposition is not unique. Given a particular mixed density operator, there is an infinite number of ways to decompose it into pure states. This is not necessarily bad because mixed states provide a more general way to represent uncertainty.

3. Alternative interpretations

3.1. Quantum-cognition system

We have presented one view of the mapping of psychological concepts into quantum concepts. Let us briefly summarize

them more formally. First we choose the dimension N of the Hilbert space. The stimulus that is presented to the participant (e.g., stimulus f) along with other experimental information prepares the state $|\psi_f\rangle$. A specific question (is the face to be presented a 'bad guy?', is it 'handsome?', 'does the person look intelligent?', 'should you attack this face?', or whatever question you wish) determines the basis for the Hilbert space. The basis vectors are used to define the projectors (e.g., D_A, D_W), and a linear combination of the projectors forms the observable (e.g. observable D). Let us call this the 'quantum cognition' system [9]. Now we consider some alternative systems that have been proposed. This comparison was inspired by working together on [29], as well as [35].

3.2. Stimulus-response system

Another approach, called the stimulus-response system, does not assume that the stimulus changes the state. Instead the state only depends on the individual $|\psi_I\rangle$ (see, e.g., [29]). The stimulus and question together define the observable $D(f)$. Consider again the C-D task in which the person sees a face and then decides an action. Before any face or other stimulus information is presented, the person is in the initial state $|\psi_I\rangle$. The face stimulus together with the decision determine the observable $D(f)$. Actually this mapping turns out not to be very different than the quantum cognition mapping. Starting from the quantum cognition system, we have

$$\begin{aligned} p(A|f) &= \|D_A \cdot |\psi_f\rangle\|^2 \\ &= \|D_A \cdot U_f^\dagger \cdot |\psi_I\rangle\|^2 \\ &= \|(U_f \cdot D_A \cdot U_f^\dagger) \cdot |\psi_I\rangle\|^2 \\ &= \|D_A(f) \cdot |\psi_I\rangle\|^2, \end{aligned}$$

which is the stimulus response system with $D_A(f) = U_f D_A \cdot U_f^\dagger$. In this way, the stimulus-response system is a generalization of quantum cognition system (equivalent when $D(f)$ is a projector); alternatively, the quantum cognition system unpacks and breaks the general function of the stimulus-response system down into its cognitive components. This difference between the two systems corresponds to the difference between the Heisenberg and Schrödinger pictures for a quantum system.

3.3. State context property system

Aerts and Gabora and colleagues apply slightly different rules, which they call the State Context Property or SCoP system [34]. They frequently work with conceptual combination problems. For example, a person may be informed that they will be considering the concept of say 'pet insect.' They are shown an example, such as a spider, and then they are asked to decide whether or not the example is a member of the concept 'pet insect.' When asked to consider a concept, like a pet, the person starts in a 'ground' state for the pet concept denoted $|\psi_p\rangle$. When asked to consider the concept of pet in the context of it being an insect 'is this a pet

insect’, the ground state is projected onto the subspace for this insect context to produce a new state $|\psi_{PI}\rangle$. The experimenter then asks a question: is this example spider a member of the category pet insect? The pet insect state is then defined as a superposition with regard to the example, i.e., superimposed about whether or not a spider is a member of pet insect. The yes, no answers to the membership question correspond to the projectors $M_Y, M_N = (I - M_Y)$, respectively, and together they determine the membership observable $M = 1 \cdot M_Y + 0 \cdot M_N$. The same membership observable M is used for all membership questions regardless of the examples that are presented. If the example is changed from spider to beetle, then the same observable M is applied to this new question.

Now we apply the SCoP system to the C-D task following [14]. We can define the ground state $|\psi_U\rangle$, as the concept ‘do I want to attack this face?’ when the category is unknown or undecided. When placed into the context ‘this is a good guy’ the state changes to $|\psi_G\rangle$, but when placed into the context ‘this is a bad guy’ the state changes instead to $|\psi_B\rangle$. The observable M represents (yes, no), which is applied to all three of these states to determine the probability of ‘attack.’ In [14], a simple three dimensional model was used to work out a SCoP model in detail. In this simple three-dimensional model, the states $|\psi_G\rangle, |\psi_B\rangle$ were designed to be orthogonal, and the unknown state $|\psi_U\rangle = \frac{1}{\sqrt{2}}(|\psi_G\rangle + |\psi_B\rangle)$ was assumed to be a superposition of the two known states.

SCoP sometimes works differently than the quantum cognition system and the stimulus response system. If we applied the quantum cognition system to the conceptual combination problem, then the person’s state $|\psi_s\rangle$ is prepared by the example ‘spider,’ which is the stimulus that is displayed to the person. The person knows that he or she is dealing with a spider, and this is not uncertain. What is uncertain is whether or not it is a pet, or more specifically, whether or not it is a pet insect. The question ‘is this a pet’ is represented by one observable M_P , and the question ‘is this an insect’ is represented by another observable M_I . Likewise, the stimulus-response system would form the observable $M(s, P)$ from the combination of the stimulus (spider) and question (is this a pet?).

3.4. Comparison

Let’s see how these systems work with a different example⁸. Quantum theory has been successfully applied to question order effects in attitude surveys [27]. Suppose a person is asked to judge whether or not a political administrator (yet to be presented) is honest and trustworthy. For half of the respondents, the picture and name of one administrator (e.g, Clinton) is shown first and a judgment is made, and this followed by the picture and name of another administrator (e.g., Gore) and another judgment is made; for the other half of the respondents, the pictures and names are shown in the opposite order. These questions form two related episodes, answers to one are relevant for answering

the other, and they are asked back to back with little or no time interval between questions, and so they can be treated as one trial. Completely unrelated types of questions are presented on other trials⁹. For these closely related type of trials with back to back measurements, large order effects are observed with all sorts of attitude questions. Consider the probability of saying ‘yes’ to the Clinton question and then ‘yes’ to the Gore question.

Using the stimulus-response approach, we define C_y as the projector for the answer ‘yes Clinton is honest and trustworthy’ to the Clinton stimulus, and we define G_y as the projector for the answer ‘yes Gore is honest and trustworthy’ to the Gore stimulus. Using the stimulus response approach we obtain the result $p(C_y, G_y) = \|G_y C_y |\psi_I\rangle\|^2$. Now suppose we define the projectors as follows $C_y = U_{CN} M_y U_{CN}^\dagger$, $G_y = U_{GN} M_y U_{GN}^\dagger$, where M_y is the generic projector for ‘yes he is honest and trustworthy’ applicable to any person, U_{CN}, U_{GN} are a unitary operators. The projector $U_{CN} M_y U_{CN}^\dagger$ represents the idea that we are examining the issue of honest and trustworthy from the Clinton perspective. Alternatively $U_{GN} M_y U_{GN}^\dagger$ represents the same question but now from the Gore perspective.

Now we show that the predictions from the quantum cognition system agree with those from the stimulus-response system for this example. The stimulus-response system expresses the probability as (see appendix for more details)

$$p(C_y, G_y) = \left\| \left(U_{GN} M_y U_{GN}^\dagger \right) \left(U_{CN} M_y U_{CN}^\dagger \right) |\psi_I\rangle \right\|^2 = \left\| M_y \left(U_{GN}^\dagger U_{CN} \right) M_y \left(U_{CN}^\dagger |\psi_I\rangle \right) \right\|^2,$$

where we made use of the length preserving property of unitary operators. The second line expresses the quantum cognition approach—the initial state (before the first stimulus) is unitarily transformed by the Clinton stimulus, then the projector for the question ‘yes he is honest and trustworthy’ is applied, and then the resulting state is unitarily transformed by the Gore stimulus, and then the projector for ‘yes he is honest and trustworthy’ is reapplied to the new state.

It remains an empirical question whether one system will ultimately provide a better representation for psychological studies as compared to another. The field is too new and we need to be in an exploratory mode. All three systems need to be investigated in a more competitive way to see which one evolves and survives to become most successful.

4. Concluding comments

Our goal for this article was to form correspondences between fundamental concepts based on quantum formalisms (such as state preparation, measurement, state evolution) and fundamental psychological concepts (such as stimulus, response, information processing). We also described several

⁸ This section is related to issues recently brought up by [35].

⁹ This is the actual procedure for some of the surveys examining order effects, see [11].

compelling examples where these concepts have been successfully applied to psychology.

It is sometimes difficult for physicists to accept the idea that the mathematical formalisms of quantum theory can be applied outside of physics. It seems to many physicists that this type of mathematics was specifically developed for quantum physical problems only. However, we disagree with this view, and instead we think that the mathematics discovered to solve problems in physics are highly suitable for solving analogous problems in psychology and even social sciences more generally [36]. The reason is that psychologists, like quantum physicists, must work with contextualized probabilistic systems that are highly sensitive to measurement, as well as ‘entangled’ systems that are strongly interconnected and difficult to decompose into separate and independent parts. We argue that the mathematical formalisms of quantum theory are highly suitable for any kind of complex system that exhibits these general kinds of properties.

5. Appendix

The purpose of this appendix is to analyze the sequence of projections $(U_{GN}M_yU_{GN}^\dagger)(U_{CN}M_yU_{CN}^\dagger)|\psi_0\rangle$. Suppose the Hilbert space is N -dimensional. We will focus on the use of three different bases for spanning this space. One is the ‘neutral’ basis $\{|N_i\rangle, i = 1, N\}$ used to represent the question ‘is the person honest and trustworthy’ for *any* person, which is used in the quantum cognition system. The other two $\{|C_i\rangle, i = 1, N\}$, $\{|G_i\rangle, i = 1, N\}$ are used to represent the Clinton/Gore questions about honest and trustworthy used in the stimulus-response system.

Define $U_{CN} = \sum_k |C_k\rangle\langle N_k|$ as the unitary operator used to change from the neutral to the Clinton basis and $U_{CN}^\dagger = \sum_k |N_k\rangle\langle C_k|$ is the inverse; $U_{GN} = \sum_k |G_k\rangle\langle N_k|$ changes from the neutral to the Gore basis and $U_{GN}^\dagger = \sum_k |N_k\rangle\langle G_k|$ is the inverse. It follows that $U_{GC} = U_{GN}U_{CN}^\dagger = (\sum_j |G_j\rangle\langle N_j|)(\sum_j |N_j\rangle\langle C_j|) = \sum_{ij} |G_i\rangle\langle N_i| \cdot |N_j\rangle\langle C_j| = \sum_k |G_k\rangle\langle C_k|$ changes to from the Clinton to the Gore basis, and $U_{GC}^\dagger = \sum_k |C_k\rangle\langle G_k|$ is the inverse.

When expressed in terms of either the neutral or Clinton basis, the $N \times N$ matrix for U_{CN} equals $V_{CN} = [\langle N_j | C_i \rangle]$, (transition to row j from column i), with $\langle N_j | C_i \rangle = \langle N_j | \sum_k |C_k\rangle\langle N_k| \cdot |N_i\rangle = \langle C_j | \sum_k |C_k\rangle\langle N_k| \cdot |N_i\rangle = \langle N_j | C_i \rangle$. Note that $V_{CN}^\dagger = [\langle C_j | N_i \rangle]$. Likewise, when expressed in terms of either the neutral or Gore basis, the $N \times N$ matrix for U_{GN} equals $V_{GN} = [\langle N_j | G_i \rangle]$, and $V_{GN}^\dagger = [\langle G_j | N_i \rangle]$. When expressed in terms of either the Clinton or Gore basis, the $N \times N$ matrix for U_{GC} equals $V_{GC} = V_{CN}^\dagger V_{GN} = [\langle C_j | G_i \rangle]$, and note that the matrix for U_{GC}^\dagger equals $V_{GC}^\dagger = V_{GN}^\dagger V_{CN} = [\langle G_j | C_i \rangle]$.

Define the initial state, at the trial beginning and before any stimulus is presented, as $|\psi_0\rangle = \sum_i |N_i\rangle\langle N_i| \psi_0\rangle = \sum_i |C_i\rangle\langle C_i| \psi_0\rangle = \sum_i |G_i\rangle\langle G_i| \psi_0\rangle$ and $\|\psi_0\|^2 = 1$. When

expressed in terms of the neutral basis, the $N \times 1$ matrix of coordinates equals $V_N(0) = [\langle N_i | \psi_0 \rangle]$, when expressed in terms of the Clinton basis, the $N \times 1$ matrix of coordinates equals $V_C(0) = [\langle C_i | \psi_0 \rangle]$, and when expressed in terms of the Gore basis, the $N \times 1$ matrix of coordinates equals $V_G(0) = [\langle G_i | \psi_0 \rangle]$.

Finally, we define the projector for the answer ‘yes’ to the question ‘is the person honest and trustworthy’ for any person. This is defined in terms of the neutral basis as $M_y = \sum_{j \in \text{yes}} |N_j\rangle\langle N_j|$. When expressed in terms of the neutral basis, this equals an $N \times N$ indicator matrix $W_y = [\langle N_j | M_y | N_i \rangle]$, with ones on the diagonal corresponding to yes, and zero otherwise.

We start by analyzing the first the Clinton question $(U_{CN}M_yU_{CN}^\dagger)|\psi_0\rangle$. First, we obtain $U_{CN}^\dagger|\psi_0\rangle = \sum_k |N_k\rangle\langle C_k| \cdot \sum_j |N_j\rangle\langle N_j| \psi_0\rangle = \sum_k |N_k\rangle\langle C_k| \sum_j \langle C_k | N_j \rangle \langle N_j | \psi_0 \rangle$. This corresponds to the matrix product $V_{CN}^\dagger \cdot V_N(0)$. Second, we apply the measurement to obtain $M_y U_{CN}^\dagger |\psi_0\rangle = \sum_{k \in \text{yes}} |N_k\rangle\langle N_k| \cdot \sum_j |N_j\rangle\langle N_j| \sum_i \langle C_j | N_i \rangle \langle N_i | \psi_0 \rangle = \sum_{k \in \text{yes}} |N_k\rangle\langle N_k| \sum_j \langle C_k | N_j \rangle \langle N_j | \psi_0 \rangle$, which corresponds to the matrix $W_y \cdot V_{CN}^\dagger \cdot V_N(0)$. Third, we apply the last unitary operator to obtain the projection $|\psi_1\rangle = (U_{CN}M_yU_{CN}^\dagger)|\psi_0\rangle = \sum_k |C_k\rangle\langle N_k| \cdot \sum_{j \in \text{yes}} |N_j\rangle\langle N_j| \sum_i \langle C_i | N_k \rangle \langle N_i | \psi_0 \rangle = \sum_{k \in \text{yes}} |C_k\rangle\langle N_k| \sum_j \langle C_k | N_j \rangle \langle N_j | \psi_0 \rangle$. This corresponds to the matrix product $V_N(1) = (V_{CN} \cdot W_y \cdot V_{CN}^\dagger) \cdot V_N(0)$. Using the fact that $\langle C_j | \psi_0 \rangle = \sum_j \langle C_k | N_j \rangle \langle N_j | \psi_0 \rangle$ the result of the first measurement can be expressed as $|\psi_1\rangle = \sum_{k \in \text{yes}} |C_k\rangle\langle C_j | \psi_0 \rangle$. The probability of yes to the first question equals the squared length $\|\psi_1\|^2 = V_N^\dagger(1) V_N(1) \leq \|\psi_0\|^2 = V_N^\dagger(0) V_N(0) = 1$. (Notice that we have not normalized this projection).

Hereafter, we operate on the projection $|\psi_1\rangle = \sum_{k \in \text{yes}} |C_k\rangle\langle C_k | \psi_0 \rangle = \sum_{k \in \text{yes}} |C_k\rangle\langle N_k| \sum_j \langle C_k | N_j \rangle \langle N_j | \psi_0 \rangle$. This projection can be also be expressed in the Gore basis as $|\psi_1\rangle = \sum_k |G_k\rangle\langle G_k | \psi_1 \rangle$. It’s matrix representation equals $V_N(1) = (V_{CN} \cdot W_y \cdot V_{CN}^\dagger) \cdot V_N(0)$.

Now we repeat these operations for the Gore question $(U_{GN}M_yU_{GN}^\dagger)|\psi_1\rangle$. First, we obtain $U_{GN}^\dagger|\psi_1\rangle = \sum_k |N_k\rangle\langle G_k| \cdot \sum_j |G_j\rangle\langle G_j| \psi_1\rangle = \sum_k |N_k\rangle\langle G_k| \cdot \sum_k \langle G_k | \psi_1 \rangle$. This corresponds to the matrix product $V_{GN}^\dagger \cdot V_N(1) = V_{GN}^\dagger \cdot (V_{CN} \cdot W_y \cdot V_{CN}^\dagger) \cdot V_N(0)$. Then we apply the measurement to obtain $M_y U_{GN}^\dagger |\psi_1\rangle = \sum_{k \in \text{yes}} |N_k\rangle\langle N_k| \cdot \sum_j |N_j\rangle\langle G_j| \psi_1\rangle = \sum_{k \in \text{yes}} |N_k\rangle\langle G_k | \psi_1 \rangle$ which corresponds to the matrix $W_y \cdot V_{GN}^\dagger \cdot V_{CN} \cdot W_y \cdot V_{CN}^\dagger \cdot V_N(0)$. Then we apply the last unitary operator to obtain the second projection $|\psi_2\rangle = (U_{GN}M_yU_{GN}^\dagger)|\psi_1\rangle = \sum_k |G_k\rangle\langle N_k| \cdot \sum_{j \in \text{yes}} |N_j\rangle\langle G_j | \psi_1 \rangle = \sum_{k \in \text{yes}} |G_k\rangle\langle G_k | \psi_1 \rangle$. This corresponds to the matrix product $V_N(2) = (V_{GN} \cdot W_y \cdot V_{GN}^\dagger)(V_{CN} \cdot W_y \cdot V_{CN}^\dagger) \cdot V_N(0)$. The final probability of yes to the first and then the second question equals the squared length of the second projection $\|\psi_2\|^2 = V_N^\dagger(2) V_N(2) \leq \|\psi_1\|^2 = V_N^\dagger(1) V_N(1) \leq \|\psi_0\|^2 = V_N^\dagger(0) V_N(0) = 1$.

In sum, using the convenient matrix representation, we find that the answer from the stimulus response system equals $\|(V_{GN} \cdot W_y \cdot V_{GN}^\dagger)(V_{CN} \cdot W_y \cdot V_{CN}^\dagger) \cdot V_N(0)\|^2 = \|W_y \cdot V_{GC}^\dagger \cdot W_y \cdot V_{CN}^\dagger \cdot V_N(0)\|^2$, where the effect of V_{GN} in the last step is ignored because it does not change length. The latter expression describes the same answer in terms of the quantum cognition system: first, the initial state $V_N(0)$, before any stimulus, is changed by a unitary matrix to the state $(V_{CN}^\dagger \cdot V_N(0))$ based on the information provided by the stimulus. Second, a measurement is taken using the general measurement matrix W_y that does not depend on any stimulus. Third the state is changed again by the unitary matrix V_{GC}^\dagger that depends on the second stimulus. Finally, the second measurement is taken using the same general measurement matrix.

References

- [1] Aerts D and Aerts S 1994 Applications of quantum statistics psychological studies of decision processes *Found. Phys.* **1** 85–97
- [2] Bordely R F 1998 Quantum mechanical and human violations of compound probability principles: toward a generalized Heisenberg uncertainty principle *Oper. Res.* **46** 923–26
- [3] Khrennikov A Y 1999 Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social, and anomalous phenomena *Found. Phys.* **29** 1065–98
- [4] Atmanspacher H, Romer H and Walach H 2002 Weak quantum theory: complementarity in physics and beyond *Found. Phys.* **32** 379–406
- [5] Penrose R 1989 *The Emperor's New Mind* (Oxford: Oxford University Press)
- [6] Penrose R 1994 *Shadows of the Mind* (Oxford: Oxford University Press)
- [7] Hameroff S 1994 Quantum coherence in microtubules. A neural basis for emergent consciousness? *J. Consciousness Stud.* **1** 91–118
- [8] Khrennikov A 2010 *Ubiquitous Quantum Structure: From Psychology to Finance* (Berlin: Springer)
- [9] Busemeyer J R and Bruza P D 2012 *Quantum Models of Cognition and Decision* (Cambridge: Cambridge University Press)
- [10] Pothos E M and Busemeyer J R 2013 Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* **36** 255–327
- [11] Wang Z, Busemeyer J R, Atmanspacher H and Pothos E M 2013 The potential of using quantum theory to build models of cognition *Top. Cogn. Sci.* **5** 689–710
- [12] Pothos E M and Busemeyer J R 2009 A quantum probability explanation for violations of rational decision making *Proc. R. Soc. B* **276** 2171–78
- [13] Busemeyer J R, Pothos E M, Franco R and Trueblood J S 2011 A quantum theoretical explanation for probability judgment errors *Psychological Rev.* **118** 193–218
- [14] Aerts D 2009 Quantum structure in cognition *J. Math. Psychol.* **53** 314–48
- [15] Gabora L and Aerts D 2002 Contextualizing concept using a mathematical generalization of the quantum formalism *J. Exp. Theor. Artif. Intell.* **14** 327–58
- [16] Blutner R, Pothos E M and Bruza P 2013 A quantum probability perspective on borderline vagueness *Top. Cogn. Sci.* **5** 711–36
- [17] Trueblood J S and Busemeyer J R 2011 A quantum probability model for order effects on inference *Cogn. Sci.* **35** 1518–52
- [18] Trueblood J S and Busemeyer J R 2012 A quantum probability model of causal reasoning *Frontiers Cogn.* **2** 289
- [19] Busemeyer J R, Wang Z and Mogiliansky-Lambert A 2009 Comparison of quantum and Markov models of decision making *J. Math. Psychol.* **53** 423–33
- [20] Wang Z and Busemeyer J R 2013 A quantum question order model supported by empirical tests of an *a priori* and precise prediction *Top. Cogn. Sci.* **5** 689–710
- [21] Conte E et al 2009 Mental states follow quantum mechanics during perception and cognition of ambiguous figures *Open Syst. Inf. Dyn.* **16** 1–17
- [22] Lambert-Mogiliansky A, Zamir S and Zwirn H 2009 Type indeterminacy: a model of the K-T (Kahneman-Tversky)-man *J. Math. Psychol.* **53** 349–61
- [23] la Mura P 2009 Projective expected utility *J. Math. Psychol.* **53** 408–41
- [24] Yukalov S and Sornette D 2011 Decision theory with prospect interference and entanglement *Theory Decis.* **70** 283–328
- [25] Townsend J T, Silva K M, Spencer-Smith J and Wenger M 2000 Exploring the relations between categorization and decision making with regard to realistic face stimuli *Pragmat. Cogn.* **8** 83–105
- [26] Busemeyer J R, Wang Z and Townsend J T 2006 Quantum dynamics of human decision making *J. Math. Psychol.* **50** 220241
- [27] Wang Z, Solloway T, Shiffrin R M and Busemeyer J R 2014 Context effects produced by question orders reveal quantum nature of human judgments *Proc. Natl. Acad. Sci. USA* **111** 9431–36
- [28] Peres A 1998 *Quantum Theory: Concepts and Methods* (Kluwer: Academic)
- [29] Khrennikov A, Basieva I, Dzhaifarov E N and Busemeyer J R 2014 Quantum models for psychological measurements: an unsolved problem (arXiv:1403.3654v1)
- [30] Miyadera T and Phillips T 2012 A quantum probability-theoretic account of human judgment errors: an axiomatic approach *Proc. 34th Annual Conf. Cogn. Sci. Soc. (Sapporo, Japan, 1–4 August 2012)* ed N Miyake, D Peebles and R P Cooper, pp 2014–8
- [31] Asano M, Basieva I, Khrennikov A, Ohya M and Yamato I 2013 Non-Kolmogorovian approach to the context-dependent systems breaking the classical probability law *Found. Phys.* **43** 895–911
- [32] Khrennikov A 2009 A model of quantum-like decision-making with applications to psychology and cognitive science *Biosystems* **95** 179–87
- [33] Camparo J 2013 A geometrical approach to the ordinal data of Likert scaling and attitude measurements: the density matrix in psychology *J. Math. Psychol.* **57** 29–42
- [34] Aerts D and Gabora L 2005 A theory of concepts and their combinations:II. A Hilbert space representation *Kybernetika* **34** 192–221
- [35] Dzhaifarov E 2014 Quantum states and observables in psychological measurements (arXiv:1405.5389v1)
- [36] Haven E and Khrennikov A 2013 *Quantum Social Science* (Cambridge: Cambridge University Press)