



Topics in Cognitive Sciences (2013) 1–22

Copyright © 2013 Cognitive Science Society, Inc. All rights reserved.

ISSN:1756-8757 print/1756-8765 online

DOI: 10.1111/tops.12040

A Quantum Question Order Model Supported by Empirical Tests of an *A Priori* and Precise Prediction

Zheng Wang,^a Jerome R. Busemeyer^b

^a*School of Communication, Center for Cognitive and Brain Sciences, The Ohio State University*

^b*Department of Psychological and Brain Sciences, Indiana University*

Received 10 May 2012; received in revised form 14 August 2012; accepted 18 August 2012

Abstract

Question order effects are commonly observed in self-report measures of judgment and attitude. This article develops a quantum question order model (the QQ model) to account for four types of question order effects observed in literature. First, the postulates of the QQ model are presented. Second, an *a priori*, parameter-free, and precise prediction, called the QQ equality, is derived from these mathematical principles, and six empirical data sets are used to test the prediction. Third, a new index is derived from the model to measure similarity between questions. Fourth, we show that in contrast to the QQ model, Bayesian and Markov models do not generally satisfy the QQ equality and thus cannot account for the reported empirical data that support this equality. Finally, we describe the conditions under which order effects are predicted to occur, and we review a broader range of findings that are encompassed by these very same quantum principles. We conclude that quantum probability theory, initially invented to explain order effects on measurements in physics, appears to be a powerful natural explanation for order effects of self-report measures in social and behavioral sciences, too.

Keywords: Order effects; Question order; Judgment; Quantum probability; Bayesian; Markov; Constructive; Vector feature space

1. Introduction

Early in the 20th century, physicists developed quantum mechanics to explain phenomena that seemed paradoxical according to classical Newtonian theory. In the process, they also discovered an entirely new theory of probability (Suppes, 1961). One of the paradoxes explained by quantum theory was the finding that the order of measurements affected the final observed probabilities. In the terminology of quantum physics, some

Correspondence should be sent to Zheng Wang, School of Communication, Center for Cognitive and Brain Sciences, The Ohio State University, Columbus, OH 43210. E-mail: wang.1243@osu.edu

observables were found to be *non-commutative*. Consequently, as the mathematical foundation for quantum mechanics, quantum probability theory was founded on a non-commutative algebra (Von Neumann, 1932). This is one of the fundamental differences from the classical probability theory (the mathematical foundation for statistical mechanics in classical Newtonian theory).

Order effects of measurements are not unique to physics. The order in which questions are asked has long been shown to influence human judgments in social and behavioral research (e.g., Sudman & Bradburn, 1974; Schuman & Presser, 1981; Tourangeau, Rips, & Rasinski, 2000). A natural inquiry is given in the following: Can quantum probability theory be usefully applied to predict human judgments and self-report measurements in social and behavioral sciences? In fact, Niels Bohr (1958), who received the Nobel Prize in 1922 for his contributions to understanding atomic structure and quantum mechanics, often speculated that ideas from quantum theory have applications outside of physics.¹

A skeptic would argue that it is quite a speculative leap to think that quantum probability theory can be applied to human judgments simply because human judgments display order effects. Indeed, to rigorously test the idea, we need to derive a quantitative prediction about order effects from quantum probability theory that can be empirically tested on human judgments and can be used to differentiate the quantum versus classical theories. This is exactly the purpose of this article. First, the postulates of the quantum question order model (the QQ model) are presented. Second, an *a priori*, parameter-free, and precise prediction, called the QQ equality, is derived from these mathematical principles, and six empirical data sets are used to test the prediction. Third, a new index is derived from the model to measure similarity between questions. Fourth, we show that in contrast to the QQ model, Bayesian and Markov models do not generally satisfy the QQ equality and thus cannot account for the reported empirical data which support this equality. Finally, we describe the conditions under which order effects are predicted to occur. Also, we review a broader range of findings that are encompassed by these very same quantum principles. We conclude that quantum probability theory, initially invented to explain order effects on measurements in physics, appears to be a powerful natural explanation for order effects of self-report measures in social and behavioral sciences, too.

2. Quantum probability theory applied to human judgments

Quantum probability theory is defined by a set of axioms for probabilities proposed by Von Neumann (1932). It is a generalization of classical probability theory, which is defined by a different set of axioms proposed by Kolmogorov (1933). Both classical and quantum probability theories are designed to assign probabilities to possible events in the world. The key difference between them lies in the mathematical nature of the definition of an event. Classical (Kolmogorov) theory defines events as subsets of a sample space; in contrast, quantum (Von Neumann) theory defines events as subspaces of a vector space. Events defined by subsets are commutative (i.e., order independent), whereas events defined by subspaces are non-commutative (i.e., order dependent). It is the

order-dependent nature of events that makes quantum theory an attractive way to represent human judgments. The social and behavioral sciences have been largely committed to the classical probability theory. While it is true that quantum probability has rarely been applied outside of physics, a growing number of researchers are exploring its usefulness for explaining human judgments. See the introduction to this special issue for an overview of quantum theory applied to cognition.²

2.1. Application to question order effects

Before introducing the QQ model, it will be helpful to have a concrete example in mind. One of the four examples presented in a review of question order effects by Moore (2002) concerned public opinions on characters of Bill Clinton and Al Gore. In a Gallup poll conducted during September 6–7, 1997, half of the 1,002 respondents were asked the following pair of questions: “Do you generally think Bill Clinton is honest and trustworthy?” and subsequently, the same question about Al Gore. The other half of respondents answered exactly the same questions but in the opposite order. Question order effects are traditionally measured by comparing the agreement rates to each object obtained in a *non-comparative* versus a *comparative* context (see Moore, 2002, for the origin of these definitions). The *non-comparative* context refers to when an object is asked first, at which point there is no specific comparison object; the *comparative context* refers to when an object is asked second, allowing a comparison with the first object. The results of the poll exhibited a striking order effect. In the non-comparative context, Clinton received a 50% agreement rate and Gore received 68%, which shows a large gap of 18%. However, in the comparative context, the agreement rate for Clinton increased to 57% while for Gore, it decreased to 60%—that is, the gap decreased to 3%. Moore (2002) defines this as a consistency effect: The difference between the objects becomes smaller in the comparative context than in the non-comparative context.

Now, we will introduce the four basic postulates of the QQ model. These four postulates are mathematically the same as the general principles of quantum theory except that the events in this application are humans’ answers to attitude questions. Quantum theory provides a geometric approach to probability theory, and so the four postulates are illustrated in a simple manner using the geometric examples shown in Figs. 1 and 2A, B. The orthogonal axes in these figures represent different possible (yes, no) answers to questions, and the vector lying within this space represents a person’s beliefs about these questions. These figures are described in more detail below, alongside the statement of each postulate. For simplicity of illustration, the figures are limited to two dimensions; but please note, the general theory is applicable to all arbitrarily high dimensional spaces (see the precise and abstract summary in Section 2.2).

Postulate 1: A person’s belief about an object in question is represented by a state vector in a multidimensional feature space (technically, an N -dimensional Hilbert space). In general, each dimension (technically, a basis vector) corresponds to a feature pattern, that is, a combination of properties that the person may or may not believe about the object. This use of feature vectors to represent belief or knowledge is consistent with

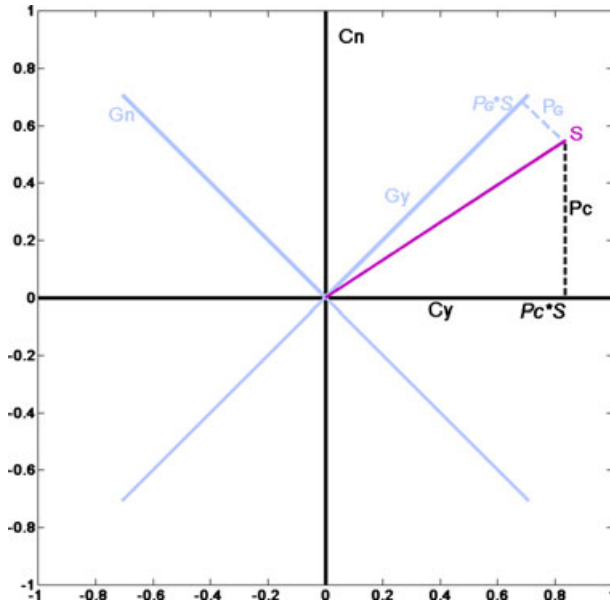


Fig. 1. A two-dimensional example of the quantum model of question order effects.

many other cognitive models of memory (e.g., Hintzman, 1988; Shiffrin & Steyvers, 1997). In our example in Fig. 1, the feature space is limited to two orthogonal dimensions for simplicity (representing either a “yes” or “no” answer to a question), and the initial belief state vector is represented by the symbol S (i.e., the purple line in Fig. 1). In addition, the belief vector is assumed to have unit length, $\|S\| = 1$, which means the squared length of the vector (i.e., sum of squared magnitudes of vector elements) equals unity. As discussed later, this is important to guarantee that the probabilities computed from the model for all possible answers to a question sum up to one.

It is important to note that the initial belief state S represents a very general state. Although it is represented by a single vector, it can be assigned coordinates with respect to different sets of basis vectors in the multidimensional feature space. In our example here in Fig. 1, the belief vector S , on one hand, can be described in terms of the Clinton basis (i.e., Cy – Cn basis vectors, black lines in Fig. 1). Then, S represents the respondent’s belief concerning “whether Clinton is honest and trustworthy” and its two coordinates are (.8367, .5477). On the other hand, S can be described using the Gore basis (i.e., Gy – Gn basis vectors, blue lines in Fig. 1). Then, it represents the respondent’s belief about “whether Gore is honest and trustworthy” and its two coordinates are (.9789, –.2043). The empirical meanings of these coordinates are explained next.

Postulate 2: A potential response to a question is represented by a subspace of the multidimensional feature space. A subspace is spanned by a subset of the basis vectors in the multidimensional feature space, which represent feature patterns describing the question at hand. In our example, the question is “Is Clinton honest and trustworthy” and one potential answer is “yes.” For our simple illustration in Fig. 1, the subspace for each

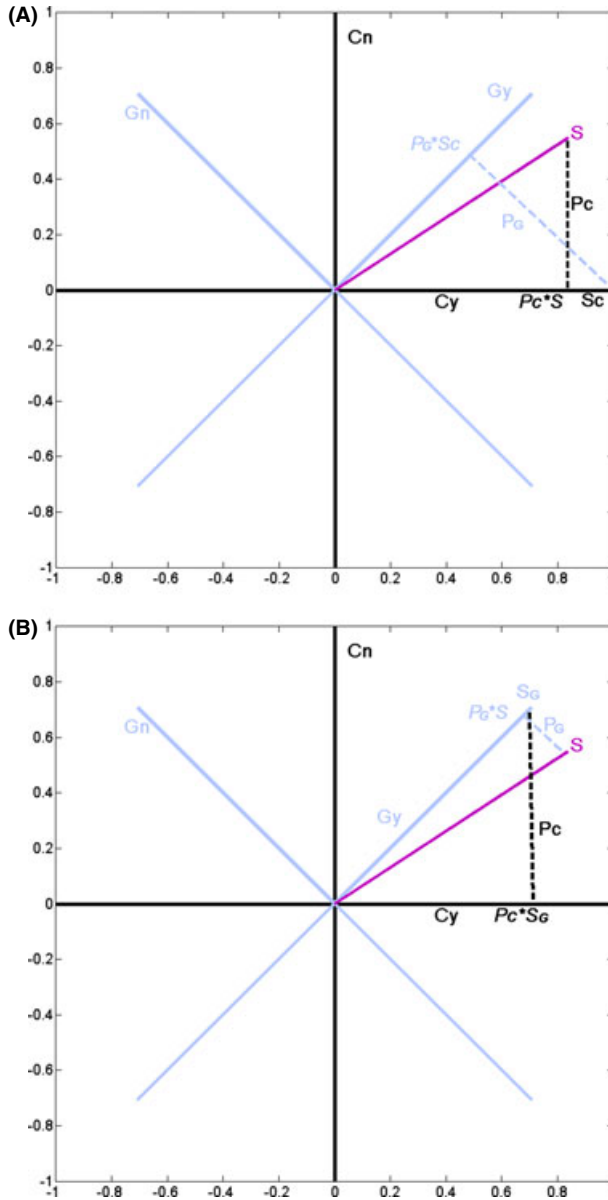


Fig. 2. (A) The probability for answering “yes” to the Gore question after having answered “yes” to the Clinton question. (B) The probability for answering “yes” to the Clinton question after having answered “yes” to the Gore question.

answer is one-dimensional, which is called a *ray*. As shown in Fig. 1, the ray labeled Cy represents responding “yes” to the Clinton question, and the ray Cn represents responding “no” to the Clinton question; similarly, Gy and Gn , respectively, represent rays of answering “yes” and “no” to the Gore question. Note that the two rays for Clinton are

orthogonal, as are the two rays for Gore. This is because the two responses—“yes” and “no”—are mutually exclusive and the probability of their conjunction should be 0. With the same reasoning, each of the Clinton rays is at an oblique (i.e., non-orthogonal) angle with respect to each of the Gore rays. This represents the fact that answering one question does not lead to a certain answer to the other question. In general, subspaces are not limited to rays. For example, a question could be represented by a plane spanned by two orthogonal basis vectors, and in this case the question is represented by the disjunction of the two properties associated with the two vectors.

In our model, responding to a question is theorized as projecting the belief state vector S down onto the subspace representing the response to that question. This is analogous to fitting the belief state to the features of the question response using multiple regression. The cognitive operation that performs this orthogonal projection is called the *projector* (technically, a Hermitian and idempotent linear operator). In our Fig. 1, the symbol P_C represents the projector for the “yes” answer to Clinton, which in our simple example is a 2×2 diagonal matrix $P_C = \text{diag}[1 \ 0]$ (when expressed in terms of the Clinton coordinates). Likewise, the symbol P_G represents the projector for the “yes” answer to Gore, which in our simple example is another 2×2 matrix denoted P_G . The projector for answering “no” to Clinton is $P_{\sim C}$, which equals $(I - P_C)$ because as explained, rays C_n and C_y are mutually exclusive (and the identity matrix I can be viewed as the projector for the entire multidimensional feature space). Similarly, the projector for “no” to Gore is $P_{\sim G}$, which equals $(I - P_G)$.

Postulate 3: The probability of responding to an opinion question equals the squared length of the projection of the state vector onto the response subspace. As mentioned, we can view the projection as the result produced by fitting the belief to the features of the question using multiple regression. Following this analogy, the squared length of the projection is similar to the proportion of the belief fit by the question response features (i.e., the R^2 obtained from a multiple regression fit). The projection of the belief S onto the C_y subspace is its projector P_C multiplied by S , that is, $P_C \cdot S$. Based on quantum probability theory, the squared length of this projection, $\|P_C \cdot S\|^2$, is the probability of responding “yes” to the Clinton question. Similarly, the probability of responding “yes” to the Gore question is $\|P_G \cdot S\|^2$. The probability of saying “no” to Clinton equals $\|P_{\sim C} \cdot S\|^2 = \|(I - P_C) \cdot S\|^2 = 1 - \|P_C \cdot S\|^2$; the probability of answering “no” to Gore is $1 - \|P_G \cdot S\|^2$. These comprise the probabilities for the Clinton and Gore questions obtained under the *non-comparative* context. In our Fig. 1 example, we know that S is represented by (0.8367, 0.5477) in the Clinton coordinates, and P_C picks the C_y coordinate of S . Hence, $p(C_y) = \|P_C \cdot S\|^2 = |0.5477|^2 = .70$. Likewise, we know that S is represented by (.9789, -.2043) in the Gore coordinates. Hence, $p(G_y) = \|P_G \cdot S\|^2 = |0.9789|^2 = .96$. Apparently in this example, when asked in a non-comparative context, Gore was favored more than Clinton.

Postulate 4: The updated belief state after deciding an answer to a question equals the normalized projection on the subspace representing the answer. This postulate concerns the change in the belief state after answering an opinion question, and it is critical in generating order effects. It is consistent with research on measurement effects on belief, attitude, intention, and behavior (Chandon, Morwitz, & Reinartz, 2005; Feldman & Lynch,

1988). This line of research has shown that survey measurement itself can change the measured evaluation and cognition. In our example, suppose the Clinton question is asked first and a respondent answers “yes,” that is, the initial belief S is being projected onto the C_y subspace. The person’s belief becomes the projection—technically, the normalized projection. Recall that $P_C \cdot S$ is the projection, but this projection usually has a length less than one and, thus, it needs to be normalized (i.e., divided by the length of the projection): $S_C = (P_C \cdot S) / \|P_C \cdot S\|$. This normalization maintains the unit length of the projection. In other words, it keeps the sum of the probabilities of all potential responses equal to one—just like the normalization used to define a conditional probability in classical probability theory. If instead, the person answers “no” to the Clinton question, then the belief becomes $S_{\sim C} = (P_{\sim C} \cdot S) / \|P_{\sim C} \cdot S\|$. This new revised state, S_C or $S_{\sim C}$, represents the change in the belief induced by expressing an opinion on Clinton and incorporates updated information consistent with the expressed opinion.

Implications of postulates for question order effects. The subsequent question must be evaluated based upon the updated belief, which has been changed by the answer to the preceding question. After updating the belief state, the process to compute probabilities of opinion responses to the next question is exactly as outlined above in the non-comparative context. For example, given that the person already said “yes” to the Clinton question and, thus, the belief has been updated to S_C , the conditional probability of saying “yes” to the Gore question equals $\|P_G \cdot S_C\|^2$ (see Fig. 2A). In the same way, the conditional probability of saying “yes” to the Clinton question, given that the person already said “yes” to the Gore question, equals $\|P_C \cdot S_G\|^2$ (see Fig. 2B).

In Fig. 2A–B, both conditional probabilities (squared length of $P_G \cdot S_C$ in Fig. 2A and of $P_C \cdot S_G$ in Fig. 2B) equal .50. This is easy to tell from the figures because the C_y and G_y rays share an angle of 45° . Now, we can calculate all the probabilities needed for the second question that arise in the *comparative* context. Consider the total probability of saying “yes” to Clinton after answering the Gore question, which we denote as TP_C . This equals the probability of saying “yes” to Gore and then “yes” to Clinton, denoted $p(GyCy)$, plus the probability of saying “no” to Gore and then “yes” to Clinton, denoted $p(GnCy)$. According to the QQ model, $p(GyCy) = \|P_G \cdot S\|^2 \cdot \|P_C \cdot S_G\|^2$ and $p(GnCy) = \|P_{\sim G} \cdot S\|^2 \cdot \|P_C \cdot S_{\sim G}\|^2$. Thus, using our example, $TP_C = p(GyCy) + p(GnCy) = (.96)(.50) + (.04)(.50) = .50$. Following the same way, the total probability of saying “yes” to the Gore question after answering the Clinton question is: $TP_G = p(CyGy) + p(CnGy) = \|P_C \cdot S\|^2 \cdot \|P_G \cdot S_C\|^2 + \|P_{\sim C} \cdot S\|^2 \cdot \|P_G \cdot S_{\sim C}\|^2 = (.70)(.50) + (.30)(.50) = .50$. Therefore, the difference in agreement rate between Clinton and Gore disappears in the comparative context.

As the simplified two-dimensional model shown, there is a large difference between the agreement rates to the two politicians in the non-comparative context, .70 for Clinton and .96 for Gore; however, there is no difference in the comparative context, .50 for both. As we intentionally designed, this example reproduces the pattern of the consistency effect as reported by Moore (2002). Observing the effect represented in Fig. 1, it is easy to tell that the consistency order effect is a result of the geometry of the feature space—particularly, the positions of the initial belief state and the two “yes” response subspaces. First, the initial belief state S starts out closer to the subspace for saying “yes” to Gore

than to Clinton. This represents the initial favoring of Gore in the non-comparative context. However, in the comparative context, the updated belief state is equally distant from these two rays, which causes the agreement rates to equalize in the comparative context (see S_C in Fig. 2A and S_G in 2B).

2.2. A summary of the QQ model

The QQ model now can be summarized as follows. A person's belief about an object is represented by a state vector, denoted S , in an N -dimensional Hilbert space. Each potential response x to a single question Q corresponds to a subspace, Qx , that has a unique orthogonal projector P_{Qx} . If responses x, y are mutually exclusive, then $P_{Qx} \cdot P_{Qy} = 0$. The projectors to all responses to a question sum to the identity $\sum_x P_{Qx} = I$. The probability of responding to a question equals the squared length of this projection, $\|P_{Qx} \cdot S\|^2$, and the updated belief state after deciding on an answer to this question equals the normalized projection $S_{Qx} = (P_{Qx} \cdot S) / \|P_{Qx} \cdot S\|$. On the basis of these assumptions, the probability of responding x to question A and then y to question B equals $p(AxBy) = \|P_{Ax} \cdot S\|^2 \cdot \|P_{By} \cdot S_{Ax}\|^2 = \|P_{By} \cdot P_{Ax} \cdot S\|^2$.

3. Quantum analysis of question order effects

As mentioned, the QQ model is not limited to a two-dimensional feature space. Hereafter, we do not assume a two-dimensional space. Instead, the following developments allow for an arbitrary N -dimensional space. An *order effect for the Clinton question* (C_C) occurs when the probability of answering "yes" to Clinton in the comparative context, given by $TP_C = p(GyCy) + p(GnCy) = \|P_C \cdot S\|^2 \cdot \|P_G \cdot S_C\|^2 + \|P_{\sim G} \cdot S\|^2 \cdot \|P_C \cdot S_{\sim G}\|^2$ differs from that in the non-comparative context, given by $p(Cy) = \|P_C \cdot S\|^2$. This difference is called an interference effect, which is defined in quantum theory as a violation of *the law of total probability* obeyed by the classical probability theory. This difference can be expressed as follows (see the Appendix for proof):

$$C_C = TP_C - p(Cy) = 2 \cdot p(GyCy) - 2 \cdot \theta \cdot \sqrt{p(Cy)} \cdot \sqrt{p(Gy)}. \quad (1a)$$

Similarly, the *order effect for the Gore question* (C_G) equals:

$$C_G = TP_G - p(Gy) = 2 \cdot p(CyGy) - 2 \cdot \theta \cdot \sqrt{p(Gy)} \cdot \sqrt{p(Cy)} \quad (1b)$$

Observing the equations and using Eq. 1a as an example, we can see that on the right-hand side of the equation, the first part, $p(GyCy)$, actually is the probability of answering "yes" to both questions in the Gore–Clinton sequence in the comparative context. The second part, $\theta \cdot \sqrt{p(Cy)} \cdot \sqrt{p(Gy)}$, is the product of three components in the non-comparative context: The magnitude of the projection onto the Cy subspaces (i.e., square root of the probability of saying "yes" to Clinton), and the magnitude of the projection onto the Gy

subspace (i.e., square root of the probability of saying “yes” to Gore), as well as an index θ representing the correlation between these two projections. Technically, $\theta = \text{Re}[\langle S_C | S_G \rangle]$, that is, the real part of the inner product between the two normalized projections. For example, in Fig. 1, we find that $\theta = \sqrt{.5} = .7071$, which is exactly the cosine between the Cy ray and the Gy ray, and reflects the 45° angle between these two rays.

This index θ merits some explanation. It is a theoretically appealing new measure derived from the QQ model. Being analogous to correlation measures, such as Pearson’s r , it is a measure of the similarity between the two “yes” projections. If the two projections are similar, then θ is positive; if they are dissimilar, then θ is negative; and $\theta = 0$ indicates that there is no relationship between the two “yes” projections (i.e., they are orthogonal). We can solve for θ in either Eqs. 1a or 1b to obtain

$$\theta = \frac{.5 \cdot C_C + p(GyCy)}{\sqrt{p(Cy)p(Gy)}} = \frac{.5 \cdot C_G + p(CyGy)}{\sqrt{p(Cy)p(Gy)}} \quad (2)$$

As shown, all the terms needed to compute θ are observable. They are often obtained in split-ballot sample surveys or experiments on questions orders. They can be computed from frequency data of the eight response sequences, such as “yes” to Clinton and then “yes” to Gore, and “yes” to Clinton and then “no” to Clinton, and so on. Therefore, using Eq. 2, researchers can easily estimate the similarity index θ based upon their own empirical data. In our example, when we estimate θ using the survey data from Moore (2002), it is .8397 and .8421 based on the two ways of calculation in Eq. 2. As expected, they are almost identical. Their average is $\theta_{CG} = .8409$, which, as expected, is positive. This large positive θ reflects and quantifies the general belief that Clinton and Gore are highly similar in their features, which is the explanation provided by Moore (2002) for the consistency effect.

4. An *a priori*, parameter-free, and precise test of the QQ model

Now, we return to one of our primary questions. Does the QQ model predict something new, interesting, and quantitative (i.e., in exact amount) on question order effects? Suppose we have two questions, A and B . According to quantum theory, the state vector has a normalized projection on the subspace for “yes” to question A (denoted by the column vector S_A), and the same state vector has a normalized projection on the subspace for “yes” to question B (denoted by the column vector S_B). According to quantum theory, the probability of transiting from S_A to S_B must equal the probability of transiting in the opposite direction (Peres, 1998). This equality follows from the property of inner products $|\langle S_B | S_A \rangle|^2 = |\langle S_A | S_B \rangle|^2$, where $\langle S_A | S_B \rangle$ is the inner product between the two vectors. This equality is called *the law of reciprocity* in quantum theory (Peres, 1998). Note that this symmetry property is a *challenging* prediction to make in the social and behavioral sciences for the following reason. In theories based on classical probability, if A and B are any two events, we generally expect that $p(A|B)$ is not the same as $p(B|A)$

(except under very unusual circumstances). However, in quantum theory, the law of reciprocity requires that if events A , B are each represented by a single dimension (i.e., a ray), such as the projections S_A and S_B , then $p(AB)$ must be exactly the same as $p(BA)$.³ Because the QQ model is strictly derived from general quantum principles, it must follow the law of reciprocity and make this challenging prediction. It is the first attempt to test whether this law works for human behavior.

Specifically in the QQ model, the law of reciprocity implies that the index θ in Eqs. 1a and 1b must be the same for both order effects. This is because it is based on the real part of the inner product between the projections of the belief state onto the answers for each question, and the real part of the inner product is independent of the order in which the questions are asked, $\text{Real}[\langle S_B | S_A \rangle] = \text{Real}[\langle S_A | S_B \rangle]$. (In general, the inner product is a complex number with real and imaginary parts.) Therefore, the exact same θ can be used to determine both of the Clinton and Gore order effects, or more generally, the A and B question order effects as defined by Eqs. 1a and 1b. This fact leads to a critical test for the QQ model called the QQ equality (see the Appendix for proof). In addition, the QQ equality depends on a critical assumption: The derivation is valid as long as the *only* factor that changes the context or the state for answering a question is answering its preceding question.

In the Appendix, it is proven that the QQ model must make the following prediction regarding the sequence of probabilities obtained from an ordered presentation of questions. Suppose two questions, A and B , can be answered with “yes” or “no” (or “agree” or “disagree,” “accept” or “reject,” and the like). The questions can be presented in the AB order for one group of respondents and in the BA order for a second group. Define $p(AyBn)$ as the probability of saying “yes” to A and then “no” to B in the AB order, define $p(BnAy)$ as the probability of saying “no” to B and then “yes” to A in the BA order, and so forth. In general, the model allows for order effects, such as $p(AyBn) \neq p(BnAy)$. However, it must *always* predict that the probability of changing answers between the two questions (i.e., one question has a “yes” answer and the other question has a “no” answer) must be equal to each other in the two question orders. That is, the QQ model precisely and always predicts the following QQ equality:

$$p(AyBn) + p(AnBy) = p(ByAn) + p(BnAy). \quad (\text{QQ Equality})$$

This prediction of QQ equality can be tested as follows. Define $p_{AB} = p(AyBn) + p(AnBy)$ as the probability of having different answers to the two questions in the AB order, and $p_{BA} = p(ByAn) + p(BnAy)$ as the probability of having different answers to the two questions in the BA order. Then, the model must *always* predict that the following q -test value equals zero:

$$q = p_{AB} - p_{BA} = 0. \quad (q\text{-test})$$

This parameter-free and quantitative test of the QQ model holds for arbitrarily high-dimensional spaces and any pair of projectors.

There are two ways to statistically test this prediction using sample data rather than population probabilities. The first way is to use a Z-test to directly test whether q , as defined above, significantly differs from 0. We can use the observed frequencies of each response sequence (e.g., $n(AyBn)$ and $n(AnBy)$) from the data to directly compute the test statistic q . Then, we can employ the standard test for a difference between proportions for two independent groups to test the quantum model prediction. If the model is correct, then for large sample sizes (N_{AB} for the AB order and N_{BA} for the BA order), the estimated q statistic will be approximately normally distributed with a mean equal to zero and population variance $\sigma^2 = 2 \cdot p \cdot (1-p)/N$, where $p = p_{AB} = p_{BA}$ (assuming the null hypothesis is true) and $N = (N_{AB}^{-1} + N_{BA}^{-1})^{-1}$ is the total number of observations. The deviation from the prediction (i.e., sample q value) is statistically significant if the Z-statistic exceeds the critical value of 1.96 (at $\alpha = .05$).

The second way is to compare (a) a binomial model of the four frequencies contained in the two-by-two table for each order condition with the constraint $p_{AB} = p_{BA}$ to (b) the same binomial model without this constraint using a χ^2 statistic. First, we can obtain maximum likelihood estimates of the probabilities for each response sequence based on the observed frequencies, under the constraint $p_{AB} = p_{BA}$ of the quantum model, and then without this constraint for the unconstrained model. The two models differ by one degree of freedom. Second, after obtaining maximum likelihood estimates of the response probabilities, we can compute a one degree of freedom χ^2 statistic, which equals $2 \cdot [\log \text{likelihood of unconstrained model} - \log \text{likelihood of constrained model}]$ (see Bishop, Fienberg, & Holland, 1975). The deviation from the quantum model prediction is statistically significant if the χ^2 statistic exceeds the critical value of 3.84 (at $\alpha = .05$).

The first way of testing the model provides a stronger test because it does not involve fitting all the predicted probabilities to all the observed sample proportions. Instead, it provides an *a priori* test of the QQ model prediction of the null hypothesis $q = p_{AB} - p_{BA} = 0$, which is directly tested using the sample estimated proportions. This test is parameter-free and precise. Such a strong test on theoretical models is rare in social and behavioral sciences. The second method requires fitting all the theoretical response probabilities, which makes it a *post hoc* test. However, this test is still useful if the q -test is violated. It allows us to examine whether there still remains a significant departure from the QQ model when we fit all of the theoretical probabilities to the data. This is a common approach to testing models in social and behavioral sciences. It is possible for the q -test to reject the model but the χ^2 test retains the model, because the model could still provide an adequate fit when allowed to fit the probabilities to the data. If the χ^2 test is also rejected, then this means the model cannot even be fitted to the data adequately.

5. Testing the QQ model using four types of order effects

Below, we present the test of the QQ model using six data sets (see Table 1). Four were reported by Moore (2002), who reviewed four different types of question order effects found in literature—consistency, contrast, additive, and subtractive. These data

Table 1
Testing the quantum model using four types of question order effects and two laboratory experiments

Statistic	Four Types of Question Order Effects Reviewed by Moore (2002)					Lab Experiments	
	Consistency	Contrast	Additive	Subtractive		Racial Hostility	AA Support
	Clinton-Gore	Gingrich-Dole	White-Black	Rose -Jackson	R-J (model)	White-Black	Race-Gender
N_{AB}	447 (509)*	437 (533)	459 (491)	506 (566)		116 (116)	118 (118)
$p(AyBy)$.4899	.3936	.3987	.3379	.3751	.2500	.4831
$p(AyBn)$.0447	.0824	.0174	.3241	.2556	.0603	.0424
$p(AnBy)$.1767	.3204	.1612	.0178	.0140	.2931	.1610
$p(AnBn)$.2886	.2037	.4227	.3202	.3553	.3966	.3136
N_{BA}	432 (494)	377 (483)	486 (514)	462 (494)		108 (108)	106 (106)
$p(ByAy)$.5625	.3448	.4012	.4156	.3749	.3241	.5660
$p(ByAn)$.1991	.3422	.0597	.0671	.0950	.0648	.1321
$p(BnAy)$.0255	.0637	.1379	.1234	.1746	.2130	.0283
$p(BnAn)$.2130	.2493	.4012	.3939	.3554	.3981	.2736
$P(Ay)$.5347	.4760	.4161	.6621	.6307	.3103	.5254
$TP(Ay)$.5880	.4085	.5391	.5390	.5496	.5370	.5943
$P(By)$.7616	.6870	.4609	.4827	.4699	.3889	.6981
$TP(By)$.6667	.7140	.5599	.3557	.3891	.5431	.6441
C_A	.0533	-.0675	.1230	-.1231	-.0811	.2267	.0689
C_B	-.0949	.0270	.0990	-.1270	-.0808	.1542	-.0540
θ	.8409	.6634	.7866	.7771	.4155	.5521	.8600
q	-.0031	-.0031	-.0189	.1514	0	.0757	.0430
z	-.1090	-.0896	-.7419	5.3026		1.2162	.8312
χ^2	.0119	.008	.5511	28.57**		1.4842	.6945

Notes. *The sample size in parentheses includes all respondents, and the sample size before the parentheses only includes respondents who provided a “yes” or “no” answer. The test statistics are based upon the latter. **The constrained maximum likelihood estimates, which produced a significant χ^2 difference test statistic, 28.57, when compared to the unconstrained model (i.e., the observed Rose-Jackson data in the column labeled “Subtractive”).

sets used split-ballot samples in Gallup national surveys. The other two were from our laboratory experiments. Interestingly, one of the six data sets (the “subtractive” data set reviewed by Moore) is predicted by the QQ model to fail the q -test.

In Table 1, each column presents a data set. The rows are organized in three groups. The first group includes the first 10 rows. The row labeled N_{AB} indicates the sample size for the group answering the questions A and B in the AB order; the next four rows present the probabilities of four response sequences to the two questions. Similarly, the row labeled N_{BA} shows the sample size for the group answering the questions in the BA order, and the next four rows present the probabilities of response sequences. Following these, the next seven rows compose the second group. Based on the probability data from the first group of rows, we can easily compute the probability of answering “yes” to A in the

non-comparative context, $p(Ay) = p(AyBy) + p(AyBn)$, and in the comparative context, $TP(Ay) = p(ByAy) + p(BnAy)$. Based on the definition of the order effect in Eq. 1a, we have $C_A = TP(Ay) - p(Ay)$. Similarly, we can compute $p(By)$, $TP(By)$, and C_B . Also, the index θ can be computed using Eq. 2. Finally, the last group of rows (the remaining three rows) in the table shows the model test statistics using the first method (q and z scores) and the second method (χ^2) described in the previous section.

The first data set in Table 1, labeled “Consistency,” refers to the “Do you generally think Bill Clinton/Al Gore is honest and trustworthy?” questions as discussed throughout the article. As shown, the prediction of the QQ model is surprisingly accurate: The predicted difference by the q -test is 0 and the observed difference in the survey data is only $-.0031$, which obviously are not statistically different. The χ^2 statistic is not significant either.

The second data set refers to the questions “Do you think Newt Gingrich/Bob Dole is honest and trustworthy?” collected during March 27–29, 1995. These data are presented in the column labeled “Contrast” in Table 1. In this case, within the non-comparative context, there was an initial difference favoring Dole; but within the comparative context, the difference became more exaggerated. In other words, the order effect was positive for Dole and negative for Gingrich. Once again, the prediction of the QQ model is extremely accurate. The prediction error, $q = -.0031$, is not significantly different from 0; and the χ^2 statistic is not significant either.

The third data set, showing an “Additive” effect, refers to the pair of questions on perceptions of racial hostility in the Aggregate of Racial Hostility Poll during June 27–30, 1996. In the poll, respondents were asked, “Do you think that only a few white people dislike blacks, many white people dislike blacks, or almost all white people dislike blacks?” preceding or following the same question asked about black hostility toward white. For both questions, the percent responding “all” or “many” increased from the non-comparative context to the comparative context—hence, it is called the “additive” order effect. The prediction of the QQ model is accurate again: Neither the prediction error q nor the χ^2 test is significant.

The fourth data set refers to the pair of question about baseball players Pete Rose and Shoeless Joe Jackson: “Do you think Rose/Jackson should or should not be eligible for admission to the Hall of Fame?” (July 13–14, 1999). These data are presented in the column labeled “Subtractive.” The survey shows that the favorable rate for both baseball players decreased in the comparative context. Interestingly, the QQ model is *predicted not to fit* this data set because this case violates an important assumption of the model.

Recall that our prediction for the current QQ model was based on the assumption that *only* the question order influences the question context. This assumption is violated in this Rose/Jackson data set. Respondents lacked sufficient knowledge about the baseball players in the questions, and it was necessary to provide some background information prior to each question. Indeed, the survey provided the following background information preceding the Pete Rose question: “As you may know, former Major League player Pete Rose is ineligible for baseball’s Hall of Fame due to charges that he had gambled on baseball games...” A little more information was provided for Joe Jackson preceding the question on him: “As you may know, Shoeless Joe Jackson is ineligible for baseball’s

Hall of Fame due to charges that he took money from gamblers in exchange for fixing the 1919 World Series...” Thus, the initial belief state when Rose/Jackson was asked first was affected by the information provided about the particular player. Furthermore, the context for the second question was changed not only by answering the first question but also sequentially by the additional background information on the player in the second question. Therefore, for this data set, the QQ model predicts that the index θ for the Rose-then-Jackson order and Jackson-then-Rose order would be different. From the perspective of the QQ model, actually the two orders in the data set were Rose background-Rose question-Jackson background-Jackson question versus Jackson background-Jackson question-Rose background-Rose question. (A more complicated version of the QQ model can be developed to account for this data set by incorporating effects of new information in between questions.)

As seen in Table 1, the Rose/Jackson data indeed show that the prediction error is highly statistically significant using both the q -test and the χ^2 method. Thus, the QQ model, as predicted, does not accurately account for the Rose/Jackson data set. Note, however, that the constrained maximum likelihood estimates do reproduce the subtractive order effect pattern (see Table 1), although the model predicted values differ significantly from the observed data. In other words, the model, on the one hand, is expected to fail to account for the exact Rose/Jackson data set; but on the other hand, it can reproduce the pattern of subtractive effect. The current QQ model is expected to accurately predict subtractive effects as long as only the order of questions influences the question context.

The final two data sets were collected from between-subjects laboratory experiments by the authors. They are shown in the last two columns of Table 1. One experiment was designed to replicate the aforementioned white/black racial hostility study. The findings indeed replicated the additive effect as found by Gallup in 1996, with a larger effect size. The prediction error, $q = .0757$, is not significantly different from 0. The χ^2 statistic is not significant either.

The other experiment was designed to replicate a survey field experiment conducted by Wilson, Moore, McKay, and Avery (2008). Participants were asked a pair of questions: “Do you generally favor or oppose affirmative action (AA) programs for racial minorities/women?” Both the study by Wilson et al., which used a national representative sample, and our laboratory experiment, found a consistency effect. Support for the AA programs for racial minorities and women become more similar in the comparative context. Specifically, support for racial minorities AA programs increased if the question followed the question on women, but support for women decreased if asked after racial minority. As shown by the test statistics in Table 1, the QQ model prediction is highly accurate again in this case: The prediction error q is not significant; neither is the χ^2 test.

6. Alternative classical models do not satisfy the q -test

Up to this point, we have shown that the QQ model can account for question order effects and *at the same time* satisfies the empirically supported QQ equality. One might

question whether other popular models of cognition based upon classical probability theory, such as Bayesian or Markov models, can also account for *both* question order effects *and* the QQ equality. If so, the q -test would not provide diagnostic evidence in favor of the QQ model. As shown below, these classical models generally predict violation of the QQ equality and thus fail to account for the reported empirical data which support the QQ equality.

6.1. Bayesian probability models

Bayesian probability models cannot explain question order effects without directly introducing assumptions about order, because they are based upon the fundamental axiom of commutativity. According to classical probability theory, the event “yes” to question A is a set (denoted A_y) contained in the sample space containing all events. Likewise, the event “yes” to question B is another set (denoted B_y) contained in the *same* sample space. The event defined by the conjunction “yes” to A and “yes” to B is the set intersection ($A_y \cap B_y = B_y \cap A_y$), which is commutative. These events do not need to be independent so that $p(B_y | A_y) \neq p(A_y | B_y)$, and the prior probabilities need not be equal so that $p(A_y) \neq p(B_y)$. However, these two probabilities must obey the product rule: $p(A_y) \cdot p(B_y | A_y) = p(A_y \cap B_y) = p(B_y \cap A_y) = p(B_y) \cdot p(A_y | B_y)$. The only way to produce order effects is to introduce two new events, denoted $O_1 = “A \text{ was asked before } B”$ and $O_2 = “B \text{ was asked before } A.”$ This allows one to construct a different sample space conditioned on each order event so that $p(A_y \cap B_y | O_1) \neq p(B_y \cap A_y | O_2)$. However, this formulation only re-describes the experimental findings in a *post hoc* manner, and it does not place any constraints on the probability sequences. Therefore, by introducing an ordering event, a Bayesian model can produce order effects, but then it does not generally predict the QQ equality to be upheld.

6.2. Markov models

Markov theory provides a classical approach for modeling probabilities of event sequences (Howard, 1971). Markov models assume that the probability distribution across states for the next time step only depends on the previous state; however, the *state* of a Markov model is designed to contain all of the necessary information about the past. In other words, the system’s “memory” is encoded into its state representation. In particular, a Markov model can be designed to incorporate memory of past answers to questions into its state specification to account for question order effects. However, unlike the QQ model, the transition probabilities between states, say S_A and S_B , of a Markov model are not constrained to satisfy the law of reciprocity, $p(S_A | S_B) \neq p(S_B | S_A)$. Consequently, Markov models generally do not predict the QQ equality to be upheld. (Detailed proof is available upon request.)

7. Conditions required for producing question order effects

According to the QQ model, a necessary condition for order effects is that the projectors for the answers to the two questions must be *non-commuting*. This is obvious from the basic quantum formula for the probability of answers x,y to questions A and then B : $p(A_x B_y) = \|\mathbf{P}_{B_y} \mathbf{P}_{A_x} \cdot S\|^2$. If the two projectors commute, $\mathbf{P}_{B_y} \mathbf{P}_{A_x} = \mathbf{P}_{A_x} \mathbf{P}_{B_y}$, then $p(A_x B_y) = \|\mathbf{P}_{B_y} \mathbf{P}_{A_x} \cdot S\|^2 = \|\mathbf{P}_{A_x} \mathbf{P}_{B_y} \cdot S\|^2 = p(B_y A_x)$ and there is no order effect. Therefore, the QQ model predicts order effects only when the two questions are *incompatible*. We discuss the issue of compatibility at more length in the general discussion section.

Given that the questions are incompatible, the similarity index θ determines the direction of the order effect. Recall that θ measures the similarity between a pair of questions. It is worth noting that the QQ model makes strong predictions for the similarity parameters when the questionnaire design is expanded to examine three rather than just two questions. For example, suppose that we have questions A, B , and C . If we use a two-dimensional model such as shown in Fig. 1, we can estimate the similarity parameter θ_{AB} for a pair of questions $A-B$, and the similarity parameter θ_{BC} for another pair of questions $B-C$. Given these two estimates, we can *a priori* predict the similarity parameter θ_{AC} for the pair of questions $A-C$, which can be derived from θ_{AB} and θ_{BC} . This is a strong test for the capability of the QQ model to predict both directions and magnitudes of order effects.

Clearly, the order effect is predicted to be a linearly decreasing function of the similarity index θ . If the similarity index is small, then a positive order effect is predicted, and if the similarity index is large, then a negative order effect is predicted. More specifically, *the order effect for question A* (C_A) depends on the relationship between θ and the ratio $a = \frac{p(B_y A_y)}{\sqrt{p(A_y)p(B_y)}}$ (we denote this ratio as a for brevity). In the same way, based on Eq. 1b, we infer that whether *the order effect for question B* (C_B) is positive or negative depends on whether θ is smaller or larger than the ratio $b = \frac{p(A_y B_y)}{\sqrt{p(A_y)p(B_y)}}$ (we denote this ratio as b for brevity). Table 2 summarizes how θ determines the direction of the order effects C_A

Table 2

The similarity measure θ_{AB} determines the directions of the order effect for object A and object B , and consequently produces the consistency, contrast, additive, or subtractive question order effect

Order Effects*	C_A **	C_B	Prob of “Yes” When Non-Comparative	Similarity Between the Two “Yes” Responses: $\cos(\theta_{AB})$ ***
Consistency	+	-	In favor of B , that is $p(A_y) < p(B_y)$	$a > \theta_{AB} > b$
Contrast	-	+	In favor of B , that is $p(A_y) < p(B_y)$	$a < \theta_{AB} < b$
Additive	+	+	No requirement	- or small +, that is, $\theta_{AB} < a, b$
Subtractive	-	-	No requirement	Large +, that is, $\theta_{AB} > a, b$

Notes. *The four types of order effects as defined by Moore (2002). **The order effect for A and B , as defined in Eqs. 1a and 1b. *** $a = \frac{p(B_y A_y)}{\sqrt{p(A_y)p(B_y)}}$ and $b = \frac{p(A_y B_y)}{\sqrt{p(B_y)p(A_y)}}$, as explained in the text.

and C_B , and further, combining with the response probabilities in non-comparative context, to produce the four types of order effects as defined by Moore (2002).

8. General conclusions and broader implications

Can quantum probability theory be applied to human judgments and self-report questions, and importantly, can we provide strong empirical tests of the predictions from the theory? To answer these questions, we developed a quantum probability model for order effects on human judgments and attitude questions, called the QQ model. We also presented an *a priori*, parameter-free, and precise q -test derived from a fundamental principle of quantum probability theory called the law of reciprocity. This law imposes a strong symmetry condition on the nature of order effects. As tested by the q statistic, the empirical results from six studies provided strong support for the QQ model.

8.1. When are questions incompatible?

According to the QQ model, two questions may be treated as either compatible or incompatible measurements, and order effects only occur for the latter type. Two questions are *incompatible* if the projectors for the questions are not defined by a common basis. Inability to form a common basis for different questions is expected to occur when the questions are new or unusual, so that an answer must be constructed online. However, if a person has a great deal of experience with a combination, then the person may have sufficient knowledge to form a compatible representation that can answer both questions using a common basis. (See Schulkin [2008] for further discussions of this type of cognitive adaption with experience.) Therefore, order effects are expected to occur for pairs of incompatible questions, especially uncommon pairs, which must be (partially) constructed on the spot to answer them. Our constructionist view of belief, attitude, and intention is consistent with many others (e.g., Feldman & Lynch, 1988; Schwarz, 2007). From this view, because of cognitive economy, belief, attitude, and intention are not stored in memory as properties but instead are constructed while being needed. The retrieved information for one question affects the context for the construction process and influences the subsequent response. Many researchers have investigated the effects of measurements themselves on measured cognition, such as mere measurement effects, self-generated validity theory, and reasons theory (Chandon et al., 2005; Dholakia & Morwitz, 2002; Feldman & Lynch, 1988; Fitzsimons & Morwitz, 1996; Janiszewski & Chandon, 2007). The proposed QQ model attempts to specify and formalize these theories.

8.2. Broader implications

The use of quantum probability theory to understand human judgments becomes more interesting when the same principles are applied to a wide range of different findings. The same four postulates presented here to account for question order effects have also

been applied to preferential choice (Lambert-Mogiliansky, Zamir, & Zwirn, 2009), categorization and decision interference (Busemeyer, Wang, & Lambert-Mogiliansky, 2009), probability judgment errors (Busemeyer, Pothos, Franco, & Trueblood, 2011), inference judgments (Trueblood & Busemeyer, 2011), similarity judgments (Pothos, Busemeyer, & Trueblood, 2013), and strategic decision making (Pothos, Busemeyer, 2009). Different applications of quantum probability theory also have been used to account for findings in perception (Atmanspacher, Filk, & Romer, 2004; Conte et al., 2009; Atmanspacher & Filk, 2013), conceptual combinations (Aerts, 2009; Aerts & Gabora, 2005; Aerts, Gabora, & Sozzo, 2013), human memory (Bruza, Kitto, Nelson, & McEvoy, 2009; Brainerd, Wang, & Reyna, 2013), risky decision making (Yukalov & Sornette, 2010), and personality measurement (Blutner & Hochnadel, 2009). Quantum random walk models have also been developed to account for dynamic decision process (Busemeyer, Wang, & Townsend, 2006; Fuss & Navarro, 2013). In conclusion, quantum probability theory provides a broad, coherent, and viable new approach to understanding human decision and cognition (Busemeyer & Bruza, 2012; Khrennikov, 1999; Pothos & Busemeyer, 2012).

Acknowledgments

This study was supported by the National Science Foundation [SES 0818277 and SES 1153846].

Notes

1. This article is not about a quantum mechanical physics model. We only use the mathematical principles of quantum probability. Sometimes this is referred to as a quantum structure (Aerts, 2009), quantum-like (Khrennikov, 2010), or generalized quantum (Atmanspacher, Romer, & Wallach, 2002) application of quantum theory.
2. For a recent review, see Busemeyer and Bruza (2012), the introduction to this special issue (Wang, Busemeyer, Atmanspacher, & Pothos, 2013), and an earlier special issue on quantum cognition (Bruza, Busemeyer, & Gabora, 2009). See Hughes (1989) for a general introduction to quantum theory.
3. Note that the law of reciprocity only holds for transitions between rays (i.e., one-dimensional subspaces) and not between higher dimensional subspaces. The QQ model can be arbitrarily high dimensional, and hence the event subspaces (i.e., attitude question measurements) and projectors to the subspaces (i.e., evaluations of the attitude questions) can be arbitrarily high dimensional. However, the projections (i.e., updated attitude state), such as S_A and S_B , are always rays, and therefore, the law of reciprocity applies to the QQ model although the model can be arbitrarily high dimensional.

References

- Aerts, D. (2009). Quantum structure in cognition. *Journal of Mathematical Psychology*, 53, 314–348.
- Aerts, D., & Gabora, L. (2005). A theory of concepts and their combinations II: A Hilbert space representation. *Kybernetes*, 34, 192–221.
- Aerts, D., Gabora, L., & Sozzo, S. (2013). Concepts and their dynamics: A quantumtheoretic modeling of human thought. *Topics in Cognitive Science*, 5(4).
- Atmanspacher, H., & Filk, T. (2013). The Necker-Zeno model for bistable perception. *Topics in Cognitive Science*, 5(4).
- Atmanspacher, H., Filk, T., & Romer, H. (2004). Quantum zero features of bistable perception. *Biological Cybernetics*, 90, 33–40.
- Atmanspacher, H., Romer, H., & Wallach, H. (2002). Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of Physics*, 32, 379–406. doi: 10.1023/A:1014809312397
- Bishop, Y. M., Fienberg, S. E., & Holland, P.W. (1975). *Discrete multivariate analysis: Theory and practice*. Cambridge, MA: The MIT Press.
- Blutner, R., & Hochnadel, E. (2009). Two qubits for C. G. Jung's theory of personality. *Cognitive Systems Research*, 11, 243–259.
- Bohr, N. (1958). *Atomic physics and human knowledge*. New York: Wiley.
- Brainerd, C., Wang, Z., & Reyna, V. (2013). Superposition of episodic memories: Overdistribution and quantum models. *Topics in Cognitive Science*, 5(4).
- Bruza, P., Busemeyer, J.R., & Gabora, L. (2009). Introduction to the special issue on quantum cognition. *Journal of Mathematical Psychology*, 53, 303–305.
- Bruza, P. D., Kitto, K., Nelson, D., & McEvoy, C. L. (2009). Is there something quantum-like about the human mental lexicon? *Journal of Mathematical Psychology*, 53, 362–377.
- Busemeyer, J., Wang, Z., & Lampert-Mogiliansky, A. L. (2009). Empirical comparison of Markov and quantum models of decision making. *Journal of Mathematical Psychology*, 53, 423–433.
- Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. *Journal of Mathematical Psychology*, 50, 220–241.
- Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum cognition and decision*. Cambridge, UK: Cambridge University Press.
- Busemeyer, J. R., Pothos, E., Franco, R., & Trueblood, J. (2011). A quantum theoretical explanation for probability judgment “errors.” *Psychological Review*, 118, 193–218.
- Chandon, P., Morwitz, V. G., & Reinartz, W. J. (2005). Do intentions really predict behavior? Self-generated validity effects in survey research. *Journal of Marketing*, 69 (2), 1–14.
- Conte, E., Khrennikov, A. Y., Todarello, O., Federici, A., Mendolicchio, L., & Zbilut, J. P. (2009). Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Systems and Information Dynamics*, 16, 1–17.
- Dholakia, U. M., & Morwitz, V. G. (2002). The scope and persistence of mere-measurement effects: Evidence from a field study of customer satisfaction measurement. *Journal of Consumer Research*, 29, 159–167.
- Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, 73, 421–435.
- Fitzsimons, G. J., & Morwitz, V. G. (1996). The effect of measuring intent on brand-level purchase behavior. *Journal of Consumer Research*, 23, 1–11.
- Fuss, I., & Navarro, D. (2013). Open, parallel, cooperative and competitive decision processes: A potential provenance for quantum probability decision models. *Topics in Cognitive Science*, 5(4).
- Hintzman, D. L. (1988). Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95, 528–551.
- Howard, R. A. (1971). *Dynamic probability systems*. New York: John Wiley & Sons.

- Hughes, R. I. G. (1989). *The structure and interpretation of quantum mechanics*. Cambridge, MA: Harvard University Press.
- Janiszewski, C., & Chandon, E. (2007). Transfer-appropriate processing, response fluency, and the mere measurement effect. *Journal of Marketing Research*, *44*, 309–323.
- Khrennikov, A. Y. (1999). Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social, and anomalous phenomena. *Foundations of Physics*, *29*, 1065–1098.
- Khrennikov, A. Y. (2010). *Ubiquitous quantum structure: From psychology to finance*. Berlin: Springer.
- Kolmogorov, A. N. (1933). *Foundations of the theory of probability*. New York: Chelsea Publishing.
- Lambert-Mogiliansky, A., Zamir, S., & Zwirn, H. (2009). Type indeterminacy: A model of the kt (Kahneman–Tversky)-man. *Journal of Mathematical Psychology*, *53*, 349–361.
- Moore, D. W. (2002). Measuring new types of question-order effects. *Public Opinion Quarterly*, *66*, 80–91.
- Niestegge, G. (2008). An approach to quantum mechanics via conditional probabilities. *Foundations of Physics*, *38*, 241–256.
- Peres, A. (1998). *Quantum theory: Concepts and methods*. New York: Kluwer Academic.
- Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability model explanation for violations of ‘rational’ decision theory. *Proceedings of the Royal Society, B*, *276* (1665), 2171–2178.
- Pothos, E. M., & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behavioral and Brain Sciences*, *36*, 255–274.
- Pothos, E. M., Busemeyer, J. R., & Trueblood, J. S. (2013). A quantum geometric model of similarity. *Psychological Review*, *120*(3), 679–696.
- Schulkin, J. (2008). *Cognitive adaptation: A pragmatist perspective*. Cambridge, MA: Cambridge University Press.
- Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys: Experiments on question form, wording, and content*. New York: Academic Press.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638–656.
- Shiffrin, R., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.
- Sudman, S., & Bradburn, N. M. (1974). *Response effects in surveys*. Chicago: Aldine.
- Suppes, P. (1961). Probability concepts in quantum mechanics. *Philosophy of Science*, *28*, 378–389.
- Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The psychology of survey response*. Cambridge, MA: Cambridge University Press.
- Trueblood, J. S., & Busemeyer, J. R. (2011). A quantum probability explanation for order effects on inference. *Cognitive Science*, *35*, 1518–1552.
- Von Neumann, J. (1932). *Mathematical foundations of quantum theory*. Princeton, NJ: Princeton University Press.
- Wang, Z., Busemeyer, J. R., Atmanspacher, H., & Pothos, E. M. (2013). The potential of using quantum theory to build models of cognition. *Topics in Cognitive Science*, *5*(4).
- Wilson, D. C., Moore, D. W., McKay, P. F., & Avery, D. R. (2008). Affirmative action programs for women and minorities. *Public Opinion Quarterly*, *72*, 514–522.
- Yukalov, V., & Sornette, D. (2010). Decision theory with prospect interference and entanglement. *Theory and Decision*, *70*, 283–328.

Appendix

The purpose of this appendix is to briefly introduce the basic axioms of quantum theory and then to derive the QQ equality. We use the Dirac bracket notation so that if $|S\rangle$ is vector, then $\langle S|$ corresponds to its adjoint representation and $\langle S|T\rangle$ represents the inner product between two vectors. According to quantum theory, events are represented as subspaces of a Hilbert space. Corresponding to each event A , there is an orthogonal

projector P_A . The state of a quantum system is represented by a unit length vector $|S\rangle$ within the Hilbert space. The probability of event A equals the squared length of the projection $p(A) = \|P_A |S\rangle\|^2$. If event A is observed, then the state is updated according to Lüder's rule $|S_A\rangle = P_A |S\rangle / \|P_A |S\rangle\|$.

Consider again the Clinton and Gore questions described in the text. We start by expanding the probability for answering "yes" to Clinton:

$$\begin{aligned}
\|P_C \cdot S\|^2 &= \|P_C \cdot I \cdot S\|^2 = \|P_C \cdot (P_G + (I - P_G)) \cdot S\|^2 = \|P_C \cdot P_G \cdot S + P_C \cdot (I - P_G) \cdot S\|^2 \\
&= \|P_C \cdot P_G \cdot S\|^2 + \|P_C \cdot (I - P_G) \cdot S\|^2 + \langle S | P_G \cdot P_C \cdot P_C \cdot (I - P_G) | S \rangle \\
&\quad + \langle S | (I - P_G) \cdot P_C \cdot P_C \cdot P_G | S \rangle \\
&= \|P_C \cdot P_G \cdot S\|^2 + \|P_C \cdot (I - P_G) \cdot S\|^2 + \langle S | P_G \cdot P_C \cdot P_C \cdot (I - P_G) | S \rangle \\
&\quad + \langle S | P_G \cdot P_C \cdot P_C \cdot (I - P_G) | S \rangle^* \\
&= \|P_C \cdot P_G \cdot S\|^2 + \|P_C \cdot (I - P_G) \cdot S\|^2 + 2 \cdot \text{Re}[\langle S | P_G \cdot P_C \cdot P_C \cdot (I - P_G) | S \rangle] \\
&= \|P_C \cdot P_G \cdot S\|^2 + \|P_C \cdot (I - P_G) \cdot S\|^2 + 2 \cdot \text{Re}[\langle S | P_G \cdot P_C \cdot (I - P_G) | S \rangle],
\end{aligned}$$

and the latter follows from the idempotent property of projectors. (The symbol x^* used in the above derivation refers to the complex conjugate of x .) Recall that

$$\begin{aligned}
TP_C &= \|P_G \cdot S\|^2 \cdot \|P_C \cdot S_G\|^2 + \|P_{\sim G} \cdot S\|^2 \cdot \|P_C \cdot S_{\sim G}\|^2 \\
&= \|P_C \cdot P_G \cdot S\|^2 + \|P_C \cdot (I - P_G) \cdot S\|^2.
\end{aligned}$$

Therefore, the order effect for Clinton can be expressed as

$$C_C = TP_C - \|P_C \cdot S\|^2 = 2 \cdot \text{Re}[\langle S | P_G \cdot P_C \cdot (I - P_G) | S \rangle].$$

Immediately, we see that if P_G and P_C commute so that $P_G \cdot P_C = P_C \cdot P_G$ then $P_G \cdot P_C \cdot (I - P_G) = P_C \cdot P_G \cdot (I - P_G) = 0$ and we predict NO order effect. Thus non-commuting projectors are a necessary condition for order effects. Now let us re-examine

$$\begin{aligned}
C_C &= TP_C - \|P_C \cdot S\|^2 = -2 \cdot \text{Re}[\langle S | P_G \cdot P_C \cdot (I - P_G) | S \rangle]. \\
&= -2 \cdot \text{Re}[\langle S | P_G \cdot P_C | S \rangle - \langle S | P_G \cdot P_C \cdot P_G | S \rangle] \\
&= -2 \cdot \text{Re}[\langle S | P_G \cdot P_C | S \rangle - \|P_C \cdot P_G \cdot S\|^2] \\
&= -2 \cdot \text{Re}[\langle S | P_G \cdot P_C | S \rangle] + 2 \cdot \|P_C \cdot P_G \cdot S\|^2 \\
&= 2\|P_G \cdot S\|^2 \cdot \|P_C \cdot S_G\|^2 - 2 \cdot \text{Re}[\langle S | P_G \cdot P_C | S \rangle].
\end{aligned}$$

In general, the inner product is a complex number which always can be expressed as $\langle S | P_G \cdot P_C | S \rangle = |\langle S | P_G \cdot P_C | S \rangle| [\cos(\phi) + i \sin \phi]$. The real part equals $\text{Re}[\langle S | P_G \cdot P_C | S \rangle] = |\langle S | P_G \cdot P_C | S \rangle| \cdot \cos(\phi)$. By defining the ratio

$$R = |\langle S | \mathbf{P}_G \cdot \mathbf{P}_C | S \rangle| / (\|\mathbf{P}_C \cdot S\| \cdot \|\mathbf{P}_G \cdot S\|),$$

then according to the Cauchy–Schwarz inequality, $0 \leq R \leq 1$. Finally, we can express

$$\begin{aligned} C_C &= TP_C - \|\mathbf{P}_C \cdot S\|^2 = 2 \cdot \|\mathbf{P}_C \cdot \mathbf{P}_G \cdot S\|^2 - 2 \cdot R \cos(\phi) \cdot \|\mathbf{P}_C \cdot S\| \cdot \|\mathbf{P}_G \cdot S\|, \\ &= 2 \cdot \|\mathbf{P}_G \cdot S\|^2 \cdot \|\mathbf{P}_C \cdot S_G\|^2 - 2 \cdot \theta \|\mathbf{P}_C \cdot S\| \cdot \|\mathbf{P}_G \cdot S\| \end{aligned}$$

with $\theta = R \cos(\phi)$ and $-1 \leq \theta \leq +1$, which is the similarity index referred to in the main text. Similarly, the order effect for Gore equals

$$C_G = TP_G - \|\mathbf{P}_G \cdot S\|^2 = 2 \cdot \|\mathbf{P}_G \cdot \mathbf{P}_C \cdot S\|^2 - 2 \cdot \text{Re}[\langle S | \mathbf{P}_C \cdot \mathbf{P}_G | S \rangle],$$

but $\text{Re}[\langle S | \mathbf{P}_C \cdot \mathbf{P}_G | S \rangle] = \text{Re}[\langle S | \mathbf{P}_G \cdot \mathbf{P}_C | S \rangle]$ so that

$$C_G = 2 \cdot \|\mathbf{P}_C \cdot S\|^2 \cdot \|\mathbf{P}_G \cdot S_C\|^2 - 2 \cdot \theta \cdot \|\mathbf{P}_C \cdot S\| \cdot \|\mathbf{P}_G \cdot S\|.$$

These two order effects share the same term,

$$2 \cdot \theta \|\mathbf{P}_C \cdot S\| \cdot \|\mathbf{P}_G \cdot S\|,$$

and therefore together they imply the relation

$$\begin{aligned} 0 &= (2 \times p(GyCy) - C_C) - (2 \times p(CyGy) - C_G) \\ &= [2 \cdot p(GyCy) - p(GyCy) - p(GnCy) + p(Cy)] \\ &\quad - [2 \cdot p(CyGy) - p(CyGy) - p(CnGy) + p(Gy)] \\ &= p(GyCy) - p(GnCy) + p(CyGy) + p(CyGn) \\ &\quad - p(CyGy) + p(CnGy) - p(GyCy) - p(GyCn) \\ &= [p(CyGn) + p(CnGy)] - [p(GyCn) + p(GnCy)] = q. \end{aligned}$$

Note that this property was derived independently by Niestegge (2008) for other purposes (an axiomatic analysis of quantum theory rather than a test of question order effects).