MBS Technical Report 15-03

# On Replacing "Quantum Thinking" with Counterfactual Reasoning[1]

Louis Narens
Department of Cognitive Sciences
Department of Logic and Philosophy of Science
Institute for Mathematical Behavioral Sciences
University of California, Irvine

**Abstract**

The probability theory used in quantum mechanics is currently being employed by psychologists to model the impact of context on decision. Its event space consists of closed subspaces of a Hilbert space, and its probability function sometimes violate the law of the finite additivity of probabilities. Results from the literature indicate that such a "Hilbert space probability theory" cannot be extended to standard, finitely additive, probability theory by the addition of new events with specific probabilities. This chapter presents a new kind of probability theory that shares many fundamental algebraic characteristics with Hilbert space probability theory but does extend to standard probability theory by adjoining new events with specific probabilities. The new probability theory arises from considerations about how psychological experiments are related through counterfactual reasoning.

**Key words**: quantum probability, orthomodular lattices, context, counterfactual probability

## 1 Introduction

Recently, ideas from quantum physics has filtered into cognitive psychology in order to to account for the troublesome influence context has in human cognition and decision making. Some investigators, following Busemeyer and co-workers (e.g., see Busemeyer and Bruza, 2012 for a summary) use physical quantum theory's concepts and methods to provide mathematical models for paradoxical empirical findings involving human decision making. Others (e.g., Dzhafarov, 2015; Narens, 2015) model similar findings through standard mathematical-psychological modeling methods and classical probability theory. This chapter focuses on a method of Narens (2015) that uses counterfactual logic to model context effects in psychological studies, and applies the method to context effects in questionnaire data.

In probability theory, classical logic is reflected in the boolean structure of its domain of events. Thus, generalizations of this structure can be looked at as generalizations of classical logic. For such generalizations to also be generalizations of probability theory, some of the properties of probability functions

[1]To appear in *Contextuality from Quantum Physics to Psychology,* edited by Ehtibar Dzhafarov, Ru Zhang, Scott Jordan, and Victor Cervantes.

may have to be appropriately generalized. This is what happened in quantum mechanics.

In the standard formulations of quantum mechanics (i.e., those based on Dirac, 1930, and von Neumann, 1932), probability functions are defined on lattice algebras of events that are more general than boolean algebra of events of events. In particular, a more general method of taking a "quantum union" $\Cup$ of events than the set-theoretic union $\cup$ is employed. This results in the event space generally having multiple complementation operators. One with special properties is selected for formulating a probability theory for quantum mechanics. It is called "orthocomplemenation" and denoted by $^{\perp}$. The resulting quantum probability theory generalizes the distributive law of boolean algebras of events,

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)\,.$$

Birkhoff & von Neumann (1936) generalized it to the modular law,

$$\text{if } \ B \subset A \ \text{ then } \ A \Cup (B \cap C) = (A \Cup B) \cap (A \Cup C)\,.$$

Husimi (1937) generalized it further to the orthomodular law,

$$\text{if } \ A \subset B \ \text{ then } \ B = A \Cup (A \cap B^{\perp})\,.$$

The modular law holds in finite dimensional Hilbert space but can fail in infinite dimensional Hilbert space. The orthomodular law always hold in infinite dimensional Hilbert space, and thus it is the correct event space law for a standard formulation of quantum mechanics.

Formally, the following concepts are used for describing event spaces. $\wp(X)$ stands for the power set of $X$, and $\subseteq$, $\cup$, $\cap$, $-$ and $\varnothing$, have their usual meanings in set theory.

- $\langle \mathcal{X}, \subseteq, \cup, \cap, -, X, \varnothing \rangle$ is said to be a *boolean algebra of events* if and only if $X \neq \varnothing$ and $\mathcal{X} \subseteq \wp(X)$.
- $\langle \mathcal{X}, \subseteq, \Cup, \Cap, \dot{-}, X, \varnothing \rangle$ is said to an *ortho-algebra of events* if and only if $X \neq \varnothing$, $\mathcal{X} \subseteq \wp(X)$, and
  - $(i)$ $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \Cup, \Cap, ^{\perp}, X, \varnothing \rangle$ is a *lattice*, that is, for all $A$ and $B$ in $\mathcal{X}$, $A \Cup B$ is the $\subseteq$-least upper bound of $A$ and $B$, $A \Cap B$ is the $\subseteq$-greatest lower bound of $A$ and $B$,
  - $(ii)$ $\mathfrak{X}$ is *complemented*, that is, for all $A$ and $B$ in $\mathcal{X}$, $A \Cup A^{\perp} = X$ and $A \Cap A^{\perp} = \varnothing$, and
  - $(iii)$ $\mathfrak{X}$ satisfies *DeMorgan's Laws,* that is, $\mathfrak{X}$ for all $A$ and $B$ in $\mathcal{X}$, $(A \Cup B)^{\perp} = A^{\perp} \Cap B^{\perp}$ and $(A \Cap B)^{\perp} = A^{\perp} \Cup B^{\perp}$.
- An ortho-algebra of events $\langle \mathcal{X}, \subseteq, \Cup, \Cap, ^{\perp}, X, \varnothing \rangle$ is said to be *orthomodular* if and only if for all $A$ and $B$ in $\mathcal{X}$,

$$\text{if } \ A \subset B \ \text{ then } \ B = A \Cup (A \Cap B^{\perp})\,.$$

The following is a well-known result: $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \Cup, \Cap, ^{\perp}, X, \varnothing \rangle$ is isomorphic to a boolean algebra of events if and only if it is an ortho-algebra of events and it satisfies *distributive law*, that is, for all $A$ and $B$ in $\mathcal{X}$,

$$A \Cup (B \Cap C) = (A \Cup B) \Cap (A \Cup C) .$$

It is well-known that for ortho-algebras that the distributive law implies the modular law which implies orthomodular law.

For dealing with probabilistic phenomena, the following concept of "probability function" is useful for orthomodular algebras:

Let $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \Cup, \Cap, ^{\perp}, X, \varnothing \rangle$ be an ortho-algebra. The $\mathbb{P}$ is said to be an *orthoprobability function* on $\mathfrak{X}$ if and only if the following three conditions hold:

1. $\mathbb{P}$ is a function from $\mathcal{X}$ into the closed unit interval of the reals, [0,1] such that $\mathbb{P}(X) = 1$ and $\mathbb{P}(\varnothing) = 0$.

2. $\Cup$-*additivity:* For all $A$ and $B$ in $\mathcal{X}$,

$$\text{if } \ B \subseteq A^{\perp}, \text{ then } \ \mathbb{P}(A \Cup B) = \mathbb{P}(A) + \mathbb{P}(B) .$$

3. $\subset$-*monotonicity:* For all $A$ and $B$ in $\mathcal{X}$, if $A \subset B$, then $\mathbb{B}(A) < \mathbb{B}(B)$.

Probability in quantum mechanics is base on a orthoprobability function on the orthomodular algebra $\mathfrak{X} = \langle \mathcal{X}, \subseteq, \Cup, \Cap, ^{\perp}, X, \varnothing \rangle$, where $X$ is the set of vectors of a Hilbert space $\mathfrak{H}$, $\mathcal{X}$ is the set of closed subspaces of $\mathfrak{H}$, $A \Cup B$ is the closed subspace of $\mathfrak{H}$ spanned by the closed subspaces $A$ and $B$, $\Cap = \cap$, and $^{\perp}$ is the operation of taking the orthogonal complement.

The usual definition of a finitely additive probability function for a boolean algebra of events is generalized by Conditions (1) and (2) above. Condition (3) is used to exclude situations where $A \subset B$ and $\mathbb{P}(A) = \mathbb{P}(B)$, suggesting that $A$ and $B$ differ by a nonempty set of probability 0. For boolean algebra of events, $\subset$-monotonicity is not needed for the kind of development presented here. However, it is needed in the proof of the following theorem, which plays a central role in this chapter's development.

**Theorem 1** *Suppose* $\mathbb{P}$ *is an orthoprobability function on the ortho-algebra* $\mathfrak{X}$*. Then* $\mathfrak{X}$ *is orthomodular.*

**Proof**. Chapter 6 of Narens (2015).

Philosophers and mathematicians have viewed an orthomodular algebra as a "quantum logic". It is a generalization of classical logic that has sufficient algebraic and logical structure to make it mathematically and philosophically interesting. It is well-known that each boolean algebra of events has a finitely additive probability function on it. The analogous statement cannot be said for orthomodular algebras of events: Greechie (1971) provided examples of finite orthomodular algebras of events that had no orthoprobability functions. Thus the hypothesis of Theorem 1 applies to situations that are more powerful than those captured by "quantum logic". It is easy to provide examples showing them to be less powerful than logics consisting of closed subspaces of a Hilbert space based on the operations $\Cup$, $\cap$, and $^{\perp}$.

In recent years, psychologists and economists have used ideas from quantum mechanics to model behavioral phenomena. A key part of their modeling is the use of closed subspaces of a Hilbert space as events and a orthomodular probability function. A natural question is whether in their modeling the probability theory can be "completed" by adding events to the algebra with probabilities assigned to them so that the Hilbert events extend to a boolean algebra of events with a finitely additive probability function. Mathematical theorems (e.g., Gleason, 1957) show that this cannot be done. However, such an extension is possible for some orthomodular algebras of events. For the latter, the orthomodular algebra of events with its orthoprobability function can be viewed as a subsystem of a classical boolean algebra of events with a finitely additive probability function.

This article shows how to construct such a system for a particular kind of experimental situation. The method of construction extends to a wide variety of other experimental situations.

# 2 An Experimental Psychological Paradigm

To simplify the presentation, only a single psychological paradigm involving questionnaire data is modeled. The paradigm is a between-subject paradigm involving two experiments.

Throughout this chapter, **P** denotes a *paradigm.* It consists of two nonempty disjoint sets, $Y_1$ and $Y_2$, called *experimental choice sets* or just *experiments* for short. Each of these choice sets have 4 distinct elements. Let

$$X = Y_1 \cup Y_2 \,.$$

$X$ is called the *domain* of **P**, and its elements are called *outcomes.*

For each experiment $Y$ in **P**, the same instructions involving $Y$'s outcomes are provided to each of $Y$'s subjects who is asked to choose exactly one of $Y$'s outcomes. The identity of the subjects and the chosen outcomes are recorded as data. By linking together the data from **P**'s experiments with known and hypothesized theory, the psychological scientist is able to draw conclusions about the subjects' psychological behavior.

**S** denotes the set of subjects participating in **P**'s experiments.

## 2.1 The Probability Function $\mathbb{P}$

The probability function $\mathbb{P}$ on $\wp(X)$ is defined as follows:

$$\mathbb{P}(F) = \frac{\sum_{a \in F} \#_s \mathcal{C}_s(a)}{\sum_{b \in X} \#_s \mathcal{C}(b)} \,, \tag{1}$$

where $F$ is an event in $\wp(X)$ and $\#_s\mathcal{C}_s(x)$ is the number of $s$ in $\boldsymbol{S}$ such that $s$ would have chosen $x$ in $X$ if she were put into the unique experiment of **P** where $x$ was an outcome. If $s$ participated in the experiment containing $x$ and $s$ chose $x$, then $s$ is counted in $\#_s\mathcal{C}_s(x)$. If $s$ did not participate in the experiment containing $x$ but $s$ would have chosen $x$ if she would had participated in that experiment, then $s$ is counted in $\#_s\mathcal{C}_s(x)$. An important implication of this is that $\mathbb{P}(F)$ does not depend of how the subjects in **S** are distributed across the experiments in **P**.

Note that it follows from the definition of $\mathbb{P}$ that $\mathbb{P}(Y_1) = \mathbb{P}(X) = 1$, because each subject in **S** either participates in $Y_1$ and chooses an outcome of $Y_1$ or does not participate in experiment $Y_1$ but has a choice of an outcome of $Y_1$ that she would have made if she were to participate in experiment $Y_1$. Similarly, $\mathbb{P}(Y_2) = 1$.

## 2.2 P's Theory

**P**'s theory, **T**, is assumed to be a logically consistent set of explicit statements about how various paradigm experiments and outcomes are related. The following notation is useful: For all events $A$ and $B$ in $\wp(X)$, let

> $\ll A \gg$ = the set of all subjects $s$ of **P** who would choose some element $a$ in $A$ if $s$ were put in the unique experiment $Y$ of **P** such that $a$ was an outcome of $Y$.

Let $A$ and $B$ be arbitrary elements of $\wp(X)$. For the purposes of this article, **P**'s theory will be formulate in terms relationships of the forms,

$$\ll A \gg \subseteq \ll B \gg \quad \text{or} \quad \ll A \gg = \ll B \gg .$$

Of course, this is a very severe limitation on the kinds of relationships that could be of interest among a paradigm's events, but it is all that is needed for the limited theory needed for the questionnaire example considered in this chapter. Using the $\ll \gg$ notation, it follows that

$$\mathbb{P}(A) = \frac{\ll A \gg}{\ll X \gg} .$$

## 2.3 Random approximation of $\mathbb{P}$

$\mathbb{P}$ is theoretical in nature. It can be approximated empirically as follows:

Suppose $Y_1$ and $Y_2$ each has $k$ subjects randomly assigned to it, where $k$ is a large number. For each $x$ in $X$, let

> $\#s(x)$ be the number of subjects $s$ in $\boldsymbol{S}$ that participated in the unique experiment in **P** which had $x$ as an outcome and chose $x$.

Let $\mathbb{P}'$ be the following function on $\wp(X)$: For each $F$ in $\wp(X)$,

$$\mathbb{P}'(F) = \frac{\sum_{a \in F} \#s(a)}{2k} \,. \tag{2}$$

Then when the paradigm has been run, $\mathbb{P}'(F)$ is the probability that some element $a$ of $F$ was *observed* to be selected. $\mathbb{P}(F)$ is the probability that a randomly selected subject $s$ in $\mathbf{S}$, *would select some outcome $a$ of $F$ if she were to participated in experiment of* $\mathsf{P}$ *that had $a$ as an outcome.* $\mathbb{P}'$ and $\mathbb{P}$ are related as follows: $2\mathbb{P}'$ is a good approximation for $\mathbb{P}$, in symbols,

$$2\mathbb{P}' \approx \mathbb{P} \,. \tag{3}$$

This is the case, because by the method of random selection, for each $a$ in $F$,

$$\#s(a) \approx \frac{\ll\{a\}\gg}{2} \,,$$

and thus,

$$2\mathbb{P}'(\{F\}) = \sum_{a \in F} \frac{\#s(a)}{k} \approx \sum_{a \in F} \frac{\ll\{a\}\gg}{2k} = \sum_{a \in F} \mathbb{P}(\{a\}) = \mathbb{P}(F) \,.$$

Equation 3 is useful relating $\mathbf{P}$'s theory to its data. For example, suppose one has a set of theoretical constraints about $\mathbb{P}$ that are formulated in terms linear equations or inequalities involving $\mathbb{P}$. By Equation 3, these can be tested against $\mathbf{P}$'s data by substituting $2\mathbb{P}'$ for $\mathbb{P}$. Because there is a constant ratio between $\mathbb{P}'$ and $\mathbb{P}$, the same constraints will approximately hold for $\mathbb{P}'$ if in reality they held for $\mathbb{P}$. However, now these constraints will be observable and available for empirical testing.

For some applications, it is only necessary for $\mathbb{P}$ to match the data approximately. For other ones, the theory makes assumptions about how the experiments of $\mathbf{P}$ are interrelated that presuppose some exact relationships among some of them. The concept of "event equivalence" described below is a formulation of such an exact relationship. In cases involving such exact relationships, the data $\mathcal{D}'$ for the estimation of $\mathbb{P}'$ is used to generate a new data set $\mathcal{D}$ that approximates $\mathcal{D}'$ while satisfying the theoretical assumptions involving the exact relationships.

## 2.4 Determinable events and propositions

In $\mathbf{P}$, the outcomes of $Y_1$ and $Y_2$ are distinct. In some behavioral paradigms, a "same outcome" is interpreted as occurring in different experiments, e.g., in a psychophysical experiment, physically identical stimuli occurring as choices in more than one paradigm experiment. This article's interpretation of "$a \in Y_i$ and $b \in Y_j$, $i \neq j$, are the same outcome" when they are in different experiments is that the data and underlying theory imply that

> the set of subjects $s$ that would choose $a$ if put in an experiment with $a$ as an outcome = the set of subjects $s$ that would choose $b$ if put in an experiment with $b$ as an outcome.

A similar concept holds more generally for events:

Events $F$ and $G$ in $\wp(X)$ are said to be *equivalent*, in symbols, $F \equiv G$, if and only if the subset of subjects $s$ of **S** who would choose some element $a$ of $F$ if she were put in the experiment of **P** that had $a$ as an outomce is the same as the subset of subjects of **S** who choose some element $b$ of $G$ if she were put in the experiment of **P** that had $b$ is otucome. Note that previous usage of "$a$ and $b$ are the same outcome, when $a \neq b$," is the same as saying, "$\{a\} \equiv \{b\}$."

It easily follows that $\equiv$ is an equivalence relation.

Distinct events in the same experiment cannot be equivalent, because of the assumption that each subject chooses one and only one outcome from each experiment. Thus some of the paradigm's theoretical assumptions must be used in order to show the equivalence of distinct events.

Sometimes, an $\equiv$-equivalence is suggested by the paradigm's empirical data. For example, if the subjects are appropriately randomly sampled and empirically all the subjects in experiment $Y_1$ chooses the alternative $a$ and all the subjects in experiment $Y_2$ chooses the alternative $b$, then one might want to use this data to make the theoretical assumption that $\{a\} \equiv \{b\}$.

The following concept is intended to roughly correspond to the concept of "observable" used in quantum physics.

An event $E$ in $\wp(X)$ is said to be *determinable* if and only if for each subject $s$ in **S**, either $(i)$ it can be determined from the paradigm's theory and data that $s$ is in $\ll E \gg$, or $(ii)$ it can be determined from the paradigm's theory and data that $s$ is in $\ll X - E \gg$. Let **Det** stand for the set of determinable events.

**Det** is not used here for the event domain of a probability function. It does not have the right kind of algebraic structure for this. Instead, a subset of it is employed as the event space.

The equivalence relation $\equiv$ partitions **Det** into equivalence classes. The $\bigcup$-union of each **Det**-$\equiv$ equivalence class is in that class, and thus each **Det**-$\equiv$ class has a $\subseteq$-largest element. This largest element is called a *proposition.* It is used to identify its equivalence class, and, throughout the remainder of this section, it is often harmlessly used in place of it.

**Prop** denotes the set of *propositions.*

The following theorem shows that $\langle \mathbf{Prop}, \subseteq \rangle$ satisfies the algebraic properties of a "orthomodular lattice".

**Theorem 2** *For all $A$ and $B$ in **Prop**, let $A \Cup B$ be the $\subseteq$-least upper bound of $A$ and $B$, $A \Cap B$ be the $\subseteq$-greatest lower bound of $A$ and $B$, and, $-A$ be the set-theoretic complement of $A$ with respect to $X$. Then $\mathfrak{P} = \langle \boldsymbol{Prop}, \subseteq, \Cup, \Cap, -, X, \varnothing \rangle$ is an orthomodular lattice and $\mathbb{P}$ restricted to **Prop** is an orthoprobability function on $\mathfrak{P}$.*

Given Theorem 1, the proof of Theorem 2 is straightforward: It is easy to verify that $\mathbb{P}$ is a finitely additive probability function on $\wp(X)$ such that for all

$A$ and $B$ in $\wp(X)$, if $A \subset B$ then $\mathbb{P}(A) < \mathbb{P}(B)$. It is not difficult to show that $\mathfrak{P}$ is an ortholattice. Thus by Theorem 1, $\mathfrak{P}$ is orthomodular. Then, using that $-$ is set-theoretic complementation, it easily follows that $\mathbb{P}$ is an orthoprobability function on $\mathfrak{P}$. Details are provided in Chapter 7 of Narens (2015).

# 3 Questionnaire Example

Busemeyer & Burza (2012) employ Hilbert spaces to model psychological choice situations. Some of these translate easily over to lattice of propositions modeling and vice versa. Others lose important psychological structure in the translation. The QQ equality discussed below is an example of the latter. It fails for some paradigm theories, but it must hold in the Hilbert space modeling. Thus, psychological explanations for its failure in a psychological paradigm cannot be translated into Hilbert space modeling.

In the abstract to their article, Wang, Solloway, Shiffrin, and Busemeyer (2014) provide the following summary for the empirical support for the QQ equality and its relationship to quantum probability theory:

> The hypothesis that human reasoning obeys the laws of quantum rather than classical probability has been used in recent years to explain a variety of seemingly "irrational" judgment and decision-making findings. This article provides independent evidence for this hypothesis based on an a priori prediction, called the quantum question (QQ) equality, concerning the effect of asking attitude questions successively in different orders. We empirically evaluated the predicted QQ equality using 70 national representative surveys and two laboratory experiments that manipulated question orders. Each national study contained 651–3,006 participants. The results provided strong support for the predicted QQ equality. These findings suggest that quantum probability theory, initially invented to explain noncommutativity of measurements in physics, provides a simple account for a surprising regularity regarding measurement order effects in social and behavioral science.

A lattice of propositions approach to psychological experimentation can be used to describe the QQ equality. However, it does not provide insights into into the equality and is not as elegant as Wang, et al.'s treatment.[2]

[2]Wang et al. models a subject answering questions as follows: The subject starts out in a belief state represented by a vector in Hilbert space. Question $Q_1$ is presented to her. The cognitive processing she uses in obtaining an answer to $Q_1$ produces a shift in in her belief state, which is represented in a new vector in the Hilbert space. While in this new state she answers question $Q_2$. The following is Wang et al.'s description of this process.

> Now we come to the heart of the model: context and order effects. As time passes and new information arrives, the content of short-term memory changes, and the belief vector changes accordingly. When two questions immediately follow each other, then after answering the first question, the belief vector that was used to answer the first question changes to match the answer just given.

The Hilbert space approach to probability is fundamentally different from a propositional lattice approach. Both are based on orthomodular lattices of events. However, a propositional lattice approach extends the orthomodular probability function to a boolean algebra of events so that both orthomodular probabilistic and standard probabilistic reasoning can be used. Hilbert space based probability theory cannot be so extended in a useful manner (Gleason, 1957).

Hilbert space has a rich geometric structure. This structure imposes restrictions on probabilistic phenomena captured by its event space that cannot be fully captured by a propositional lattice modeling approach. As a consequence, the Hilbert space approach provides modeling opportunities that are not available to a propositional lattice approach. Similarly, a propositional lattice approach provides modeling opportunities that capture phenomena that are outside of the Hilbert space modeling approach. An example of the latter involving the QQ equality is provided below. This suggests to me that explanations of the kinds of probabilistic psychological phenomena encountered in experimental psychology will require both approaches to probabilistic modeling and likely additional ones.

The following thought-experiment describes a paradigm **P** in which the QQ equality fails. The failure results from constraints on the relationships among **P**'s experiments. This is shown by the use of counterfactual reasoning applied to what a subject would have done if put into a different experiment. The paradigm is similar in design to the questionnaire studies used in Wang, et al. (2014).

**P** has two experiments, $Y_1$ and $Y_2$. They consist in presenting two Yes-No questions, $Q_1$ and $Q_2$. Each subject is put into $Y_1$ or $Y_2$. The subjects in $Y_1$ are presented the questions so that $Q_1$ is immediately followed by question $Q_2$, and the subjects in $Y_2$ are presented them so that $Q_2$ is immediately followed by $Q_1$. Subjects' data for $Y_1$ and $Y_2$ are coded as one of the following eight ordered pairs of answers, where, for example, $Q_{1,y}Q_{2,n}$ stands for, "In experiment $Y_1$ the subject answered 'Yes' for the first presented question, $Q_1$, described in the

In other words, the belief vector realigns with the current contents of short-term memory (which includes answers to previous questions) and the perspectives that flow from those contents. In geometric terms, the new belief vector used for the second question is simply the projection of the initial belief vector onto the subspace used to answer the first question, normalized to have unit length. This new belief vector is then projected onto the subspace used for answering the second question, and squared to produce the probability of a "yes" response to the second question. Using this process, the probability of the sequence of "yes" answers equals the squared length of the projection produced by first projecting the belief state onto the subspace for answering "yes" to the first question, and then projecting the updated belief vector onto the subspace for answering "yes" to the second question. When the two subspaces lie at oblique angles with respect to each other, the order of answering the questions will change the projections and ultimately the probabilities of the responses, and this is described as "noncommutativity" in quantum theory. This is where a context effect arises.

ordered pair, and 'No' for the second presented question, $Q_2$, described in the ordered pair":

$$Y_1: \ Q_{1,y}Q_{2,y},\ Q_{1,y}Q_{2,n},\ Q_{1,n}Q_{2,y},\ Q_{1,n}Q_{2,n}\ ,$$

$$Y_2: \ Q_{2,y}Q_{1,y},\ Q_{2,y}Q_{1,n},\ Q_{2,n}Q_{1,y},\ Q_{2,n}Q_{1,n}\ .$$

The notation

$$\ll Q_{1,y}Q_{2,y} \gg$$

stands for the set of all of **P**'s subjects $s$ such that $s$ would provide $Q_{1,y}Q_{2,y}$ as data in experiment $Y_1$. Similar notation is used for other ordered pairs of questions, e.g., $\ll Q_{2,y}Q_{2,y} \gg$, $\ll Q_{1,y}Q_{2,n} \gg$, etc.

The following completes the specification of **P**:

- **S** denotes the set of **P**'s subjects. **P**'s subjects are randomly distributed in equal numbers across experiments $Y_1$ and $Y_2$.
- $\mathbb{P}$ denotes the previously defined probability function on $\wp(X)$ (Equation 1).
- **P**'s theory consists of the following six constraints:
  1. $\ll Q_{1,y}Q_{2,y} \gg \subset \ll Q_{2,y}Q_{1,y} \gg$
  2. $\ll Q_{1,n}Q_{2,n} \gg \subset \ll Q_{2,n}Q_{1,n} \gg$
  3. $\ll Q_{2,y}Q_{1,y} \gg - \ll Q_{1,y}Q_{2,y} \gg \subseteq \ll Q_{1,n}Q_{2,y} \gg$
  4. $\ll Q_{2,n}Q_{1,n} \gg - \ll Q_{1,n}Q_{2,n} \gg \subseteq \ll Q_{1,y}Q_{2,n} \gg$
  5. $\ll Q_{1,n}Q_{2,y} \gg \subseteq \ll Q_{2,y}Q_{1,y} \gg$
  6. $\ll Q_{1,y}Q_{2,n} \gg \subseteq \ll Q_{2,n}Q_{1,n} \gg$

The intuition behind this thought-experiment is that a polarized population is being assessed though questions $Q_1$ and $Q_2$. The polarization is expressed by subjects' biases for and against the views expressed in questions $Q_1$ and $Q_2$. $Q_2$ is the more polarizing question, and constraints 1 and 2 of the theory states that if presented first it influences in the direction of responses of a subject's answer to $Q_1$ in direction of their biases for their answer to question $Q_2$. An answer to $Q_1$ is assumed to have no influence on the more polarizing question $Q_2$. The assumed influences of the questions on one another is most apparent in the constraints 3 and 4. For example, in constraint 3, consider a subject $s$ who would, if put in experiment $Y_2$, answer "Yes" to both $Q_2$ and $Q_1$, and, if put in experiment $Y_1$, would not answer "Yes" to both $Q_1$ and $Q_2$. Then, using constraint 1, it follows from this that $s$ must answer "No" to $Q_1$ if put in experiment $Y_1$. Constraint 3 assumes that $(i)$ her biases resulting from the presentation of $Q_1$ and her "No" response to $Q_1$ are not sufficient to change her leaning toward a "Yes" response to $Q_2$ if $Q_1$ were not presented first, but $(ii)$ her strong bias in favor of $Q_2$ and her "Yes" response to $Q_2$ is sufficiently strong to overcome her weak leaning toward a "No" to $Q_1$ if $Q_2$ were not presented first, resulting in a "Yes" response to $Q_2$. The impact of constraint 4 on "No" answers to questions $Q_1$ and $Q_2$ is similar to constraint's 3 impact on "Yes" answers to those questions. Constraint 5 says that subjects who would answer "No" to $Q_1$ and "Yes" to $Q_2$ in experiment $Y_1$, would, if put in experiment $Y_2$, answer "Yes" to $Q_2$ and "Yes" to $Q_1$. This is reasonable, because the "Yes" to $Q_2$ indicates, by assumption, a strong bias in favor of $Q_2$. By assumption, this bias impacts the weakly initially held belief against $Q_1$ (i.e., as evidenced by a "No" answer when $Q_1$ is the first question) by reversing the "No" response into a "Yes" response when $Q_1$ is the second answered question that follows the first answered question $Q_2$. (6) has a similar rationale.

Part of Wang, et al. conclusions about the QQ equality are based on pairs of questions from questionnaires of other investigators who were interested in issues other than testing the QQ equality. The following are two examples of the kinds of question pairs they employed:

- Is Bill Clinton honest and trustworthy? Is Al Gore honest and trustworthy?
- Do white people dislike black people? Do black people dislike white people?

For such kinds of pairs, Wang, et al. found strong support for the QQ equality. They noted that the QQ equality failed in other question pairs when they were separated by intervening questions. This is why the immediate following of one question of pair by the other is important in their data for obtaining the QQ equality. They provided no psychological theory as to why the QQ equality is plausible for the kinds of content presented in the questions. They, however, provided psychological explanations for the observed effects of the answer patterns significantly differing when questions are presented so that they immediately follow one and another and are in opposite temporal orders. They also provide psychological theory and data showing that the QQ equality fails when the two questions separated by intervening ones. They demonstrate mathematically that QQ equality follows from a general modeling method using Hilbert space probability theory applied to situations like questionnaire data and the assumption that one question of the pair immediately follows other.

The thought-experiment described by paradigm **P** suggests that having one question immediately following the other is not sufficient for QQ equality to hold. This is because paradigm **P** violates the QQ equality. In particular, it violates,

$$[\mathbb{P}(Q_{1,y}Q_{2,y}) - \mathbb{P}(Q_{2,y}Q_{1,y})] + [\mathbb{P}(Q_{1,n}Q_{2,n}) - \mathbb{P}(Q_{2,n}Q_{1,n})] = 0\,, \qquad (4)$$

a key component of the QQ equality. This following argument shows this.

By the definitions of $\ll Q_{1,y}Q_{2,y} \gg$ and $\ll Q_{1,n}Q_{2,y} \gg$,

$$\ll Q_{1,y}Q_{2,y} \gg \cap \ll Q_{1,n}Q_{2,y} \gg = \varnothing\,.$$

It follows from constraints (1) and (5) that

$$\ll Q_{1,y}Q_{2,y} \gg \cup \ll Q_{1,n}Q_{2,y} \gg \;\subseteq\; \ll Q_{2,y}Q_{1,y} \gg \;.$$

Thus,

$$\ll Q_{1,n}Q_{2,y}\gg \subseteq \ll Q_{2,y}Q_{1,y}\gg - \ll Q_{1,y}Q_{2,y}\gg .$$

By constraint 3,

$$\ll Q_{2,y}Q_{1,y}\gg - \ll Q_{1,y}Q_{2,y}\gg \subseteq \ll Q_{1,n}Q_{2,y}\gg .$$

Therefore,

$$\ll Q_{2,y}Q_{1,y}\gg - \ll Q_{1,y}Q_{2,y}\gg = \ll Q_{1,n}Q_{2,y}\gg . \tag{5}$$

Similarly, it follows from constraints 2, 4, and 6 that

$$\ll Q_{2,n}Q_{1,n}\gg - \ll Q_{1,n}Q_{2,n}\gg = \ll Q_{1,y}Q_{2,n}\gg . \tag{6}$$

Subtracting Equation 6 from Equation 5 yields,

$$\begin{aligned} &[\ll Q_{2,y}Q_{1,y}\gg - \ll Q_{1,y}Q_{2,y}\gg] - [\ll Q_{2,n}Q_{1,n}\gg - \ll Q_{1,n}Q_{2,n}\gg] \\ &= [\ll Q_{2,y}Q_{1,y}\gg - \ll Q_{1,y}Q_{2,y}\gg] + [\ll Q_{1,n}Q_{2,n}\gg - \ll Q_{2,n}Q_{1,n}\gg] \quad (7) \\ &= \ll Q_{1,n}Q_{2,y}\gg - \ll Q_{1,y}Q_{2,n}\gg . \quad (8) \end{aligned}$$

It follows from Expressions 7 and 8 that

$$\begin{aligned} &[\mathbb{P}(\ll Q_{2,y}Q_{1,y}\gg) - \mathbb{P}(\ll Q_{1,y}Q_{2,y}\gg)] \quad (9) \\ &\quad + [\mathbb{P}(\ll Q_{1,n}Q_{2,n}\gg) - \mathbb{P}(\ll Q_{2,n}Q_{1,n}\gg)] \\ &= \mathbb{P}(\ll Q_{1,n}Q_{2,y}\gg) - \mathbb{P}(\ll Q_{1,y}Q_{2,n}\gg) . \end{aligned}$$

When $\ll Q_{1,n}Q_{2,y}\gg \neq \ll Q_{1,y}Q_{2,n}\gg$, Equation 9 becomes a violation of Equation 4, that is, the QQ equality fails. For this thought-experiment, one would expect that only in the rarest circumstances would $\ll Q_{1,n}Q_{2,y}\gg = \ll Q_{1,y}Q_{2,n}\gg$.

The thought experiment suggests to me that the kind of questionnaire data used in Wang et al. (2014) relies on a special sort of interaction between the content of questions and the kind of subjects selected, and is not a result of general features of cognitive processing. It is likely that it is very difficult to specify the nature of the content of the questions that gives rise to the equality. It is not difficult to provide examples of paradigms whose theories imply the equality. The Hilbert space probability modeling of the equality has the advantage that it follows from a general modeling process for cognitive decisions—a process that was not originally designed for the QQ equality. However, as suggested by the thought-experiment, it may be too specific in its modeling method for some decision situations.

## 4 Conclusions

The incorporation of context into psychological probabilistic models is a major concern in psychology. Ideas from quantum mechanics—particularly the use of Hilbert space based probability theory—has recently shown promise in handling various contextual psychological impacts. It is natural to inquire whether or not it can be handled equally well or better through the use of traditional probabilistic concepts. The answer, I believe, is mixed. Hilbert space based probability theory provides a new form of geometrically modeling for psychology that incorporates probabilistic and dynamical concepts for the modeling of the influence of context on behavior. It has two foundational problems. Like other geometrical scaling techniques used in psychology it lacks rational justifications—other than "success"—for its use in psychological modeling, and there is a similar lack of justification for its probabilistic theory. The construction of a paradigm's event space presented in this chapter does not have these issues: Its events are defined directly in terms of experimental data and underlying psychological theory linking experiments. Its event structure and probabilities extend to a traditional probability theory on a boolean algebra of events. The existence of such an extension implies that the probabilities assigned to the paradigm's propositions is done in a rational manner, because such assignments in the extension were done in a rational manner according to well-known theories of "rationality" for traditional probability theory.

## References

Birkhoff, G. & von Neumann, J. (1936). The logic of quantum mechanics. *Annals of Mathematics, 37(4),* 823–843.

Busemeyer, J. R. & Bruza, P. D. (2012). *Quantum Models of Cognition and Decision.* Cambridge: Cambridge University Press.

Dzhafarov, E. N. & Kujala, J. V. (2015). Contextuality is about identity of random variables. To be published in *Physica Scripta.*

Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *J. Math. Mech., 6(6),* 885–893.

Greechie, R. J. (1971). Orthomodular lattices admitting no states. *Journal of Combinatorial Theory, 10,* 119-132.

Husimi, K. (1937). Studies in the foundations of Quantum Mechanics. *Proceedings of the Physico-Mathematical of Japan, 19,* 766–89.

Narens, L. (2015). *Probabilistic Lattices with Applications to Psychology.* Singapore: World Scientific.

von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik.* Berlin: Springer.

Wang, Z., Solloway, T., Shiffrin, R. M., & Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *PNAS, 111, (26)* 9431–9436.