QUANTUM MODELS OF COGNITION AND DECISION

Jerome R. Busemeyer
Indiana University

Supported by AFOSR
WHAT IS THE GOAL OF QUANTUM COGNITION AND DECISION?

- Not a physical/neurobiological theory of the brain
- Not a theory of consciousness
- It is a mathematical theory about human behavior
- Specifically judgments and decisions
ORGANIZATION OF THIS TALK

1. Why use quantum theory for cognition and decision?
2. Quantum vs classic probability theory.
3. Evidence for quantum probability theory.
4. Quantum versus Markov dynamics.
5. Evidence for quantum dynamics
6. Conclusions
I. WHY USE QUANTUM THEORY?
1. **Quantum** theory is a general **Axiomatic** theory of probability

   - Human judgments and decisions are **probabilistic**
   - These probabilities **do not** obey the Kolmogorov axioms
   - Quantum theory provides a **viable** alternative

2. **Non Commutativity** of measurements

   - Measurements **change** psychological states producing context effects
   - Principle of **complementarity** was borrowed by Niels Bohr from William James

3. **Vector space** representation of probabilities

   - Agrees with connectionist-neural network models of cognition
2. HOW DO WE USE QUANTUM THEORY?
COMPARISON OF CLASSIC AND QUANTUM PROBABILITY THEORIES

Kolmogorov

Von Neumann
Classical

• Each unique outcome is a member of a set of points called the Sample space

Quantum

• Each unique outcome is an orthonormal vector from a set that spans a Vector space
Classical

• Each unique outcome is a member of a set of points called the **Sample space**

• Each event is a subset of the sample space

Quantum

• Each unique outcome is an orthonormal vector from a set that spans a **Vector space**

• Each event is a subspace of the vector space.
**Classical**

- Each unique outcome is a member of a set of points called the **Sample space**
- Each event is a **subset** of the sample space
- **State** is a probability function, $p$, defined on subsets of the sample space.

**Quantum**

- Each unique outcome is an orthonormal vector from a set that spans a **Vector space**
- Each event is a **subspace** of the vector space.
- **State** is a unit length vector, $S$, defined by:

$$p(A) = ||P_A S||^2$$
Classical

• Suppose event A is observed (state reduction):

\[ p(B \mid A) = \frac{p(B \cap A)}{p(A)} \]

Quantum

• Suppose event A is observed (state reduction):

\[ p(B \mid A) = \frac{\left| P_B P_A S \right|^2}{\left| P_A S \right|^2} \]
Classical

• Suppose event A is observed (state reduction):

\[ p(B \mid A) = \frac{p(B \cap A)}{p(A)} \]

• Commutative Property

\[ p(B \cap A) = p(A \cap B) \]

Quantum

• Suppose event A is observed (state reduction):

\[ p(B \mid A) = \frac{||P_B P_A S||^2}{||P_A S||^2} \]

• Non-Commutative

\[ ||P_B P_A S||^2 \neq ||P_A P_B S||^2 \]
3. WHAT IS THE EMPIRICAL EVIDENCE?
CONJUNCTION - DISJUNCTION
PROBABILITY JUDGMENT ERRORS

Tversky & Kahneman
(1983, Psychological Review)

Busemeyer, Pothos, Franco, Trueblood
(2011, Psychological Review)
Read the following information:
Linda was a philosophy major as a student at UC Berkeley and she was an activist in social welfare movements.

Rate the probability of the following events

Linda is a feminist (.83)
Linda is a bank teller (.26)
Linda is a feminist and a bank teller (.36)
Linda is a feminist or a bank teller (.60)
LAW OF TOTAL PROBABILITY

\[ p(B) = p(F)p(B \mid F) + p(\neg F)p(B \mid \neg F) \]
\[ \geq p(F)p(B \mid F) \]

CONJUNCTION - FALLACY

VIOLATES THIS LAW
Quantum Model Predictions

\[ \left| \left| P_B S \right| \right|^2 = \left| \left| P_B I S \right| \right|^2 = \left| \left| P_B (P_F + P_{\bar{F}})S \right| \right|^2 \]

\[ = \left| \left| P_B P_F S + P_B P_{\bar{F}} S \right| \right|^2 \]

\[ = \left| \left| P_B P_F S \right| \right|^2 + \left| \left| P_B P_{\bar{F}} S \right| \right|^2 + \text{Int} \]

\[ \text{Int} = \left\langle S' P'_F P'_B P_{\bar{F}} S \right\rangle + \left\langle S' P_{\bar{F}} P'_B P_F S \right\rangle \]

\[ \text{Int} < - \left| \left| P_B P_{\bar{F}} S \right| \right|^2 \]
DISJUNCTION FALLACY

Finding: \( p(F) \geq p(F \text{ or } B) \)

\[
p(F) = 1 - \left| \left| P_F S \right| \right|^2
\]

\[
p(F \text{ or } B) = 1 - \left| \left| P_{\overline{F}} P_{\overline{B}} S \right| \right|^2
\]

Finding \( \rightarrow \left| \left| P_{\overline{F}} P_{\overline{B}} S \right| \right|^2 \geq \left| \left| P_F S \right| \right|^2 \)
INTERFERENCE OF CATEGORIZATION ON DECISION

Psychological version of a double slit experiment

Participants shown pictures of faces

**Categorize** as “good” guy or “bad” guy

**Decide** to act “friendly” or “aggressive”
Two Conditions:

**C-then-D:** Categorize face first and then decide action

**D-alone:** Decide without categorization
LAW OF TOTAL PROBABILITY

\[ p(A) = p(G)p(A | G) + p(B)p(A | B) \]

G = good guy, B = Bad guy, A = Attack

D alone Condition

C-then-D Condition
### RESULTS

| Face | p(G) | p(A|G) | p(B) | p(A|B) | TP  | P(A) |
|------|------|--------|------|--------|-----|------|
| Good | 0.84 | 0.35   | 0.16 | 0.52   | 0.37| 0.39 |
| Bad  | 0.17 | 0.41   | 0.82 | 0.63   | 0.59| **0.69** |
\[ p(A \mid D \text{ alone}) = \left| \left| P_A S \right| \right|^2 = \left| \left| P_A \cdot I \cdot S \right| \right|^2 \]
\[ = \left| \left| P_A \cdot (P_G + P_B) \cdot S \right| \right|^2 \]
\[ = \left| \left| P_A \cdot P_G \cdot S + P_A \cdot P_B \cdot S \right| \right|^2 \]
\[ = \left| \left| P_A \cdot P_G \cdot S \right| \right|^2 + \left| \left| P_A \cdot P_B \cdot S \right| \right|^2 + \text{Int} \]
\[ \text{Int} = \left< S \mid P_G P_A P_A P_B \mid S \right> + \left< S \mid P_B P_A P_A P_G \mid S \right> \]
\[ \text{Finding} \rightarrow \text{Int} > 0 \]
VIOLATIONS OF RATIONAL DECISION THEORY

Shafir & Tversky
(1992, Psychological Science)

Pothos & Busemeyer
(2009, Proceedings of Royal Society)
Examined three conditions in a prisoner dilemma task

**Known Coop:** Player is told other opponent will cooperate

**Known Defect:** Player is told other opponent will defect

**UnKnown:** Player is told nothing about the opponent
LAW OF TOTAL PROBABILITY

\[ p(PD) = \text{probability player defects} \]

when opponent's move is unknown

\[ p(PD) = p(OD)p(PD \mid OD) + p(OC)p(PD \mid OC) \]

Empirically we find: \( p(PD \mid OD) \geq p(PD \mid OC) \)

\[ \rightarrow p(PD \mid OD) \geq p(PD) \geq p(PD \mid OC) \]
## DEFECT RATE FOR TWO EXPERIMENTS

<table>
<thead>
<tr>
<th>Study</th>
<th>Known Defect</th>
<th>Known Coop</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shafir (1992)</td>
<td>0.97</td>
<td>0.84</td>
<td>0.63</td>
</tr>
<tr>
<td>Matthew (2006)</td>
<td>0.91</td>
<td>0.84</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Avg</strong></td>
<td><strong>0.94</strong></td>
<td><strong>0.84</strong></td>
<td><strong>0.65</strong></td>
</tr>
</tbody>
</table>
$p(PD) = \| P_{PD} S \|^2 = \| P_{PD} \cdot I \cdot S \|^2$

$= \| P_{PD} \cdot (P_{OD} + P_{OC}) \cdot S \|^2$

$= \| P_{PD} \cdot P_{OD} \cdot S + P_{PD} \cdot P_{OC} \cdot S \|^2$

$= \| P_{PD} \cdot P_{OD} \cdot S \|^2 + \| P_{PD} \cdot P_{OC} \cdot S \|^2 + \text{Int}$

$\text{Int} = \langle S | P_{OC} P_{PD} P_{PD} P_{OD} | S \rangle + \langle S | P_{OD} P_{PD} P_{PD} P_{OC} | S \rangle$
ATTITUDE QUESTION ORDER EFFECTS

Moore
(2002, Public Opinion Quarterly)

Wang, Solloway, Shiffrin, & Busemeyer
(2013, Proceedings National Academy of Science)
A Gallup Poll question in 1997, N = 1002, “Yes”

- Do you generally think Bill Clinton is honest and trustworthy? (50%)
- How about Al Gore? (60%)
- Do you generally think Al Gore is honest and trustworthy? (68%)
- How about Bill Clinton? (57%)
### Observed proportions in the two question orders

<table>
<thead>
<tr>
<th></th>
<th>Clinton-Gore</th>
<th></th>
<th>White-Black</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gy</td>
<td>Gn</td>
<td></td>
<td>By</td>
</tr>
<tr>
<td>Cy</td>
<td>.4899</td>
<td>.0447</td>
<td>Gore-Clinton</td>
<td>.3987</td>
</tr>
<tr>
<td>Cn</td>
<td>.1767</td>
<td>.2886</td>
<td></td>
<td>.1612</td>
</tr>
<tr>
<td>戈尔-克林顿</td>
<td>.5625</td>
<td>.0255</td>
<td></td>
<td>.4012</td>
</tr>
<tr>
<td></td>
<td>.1991</td>
<td>.2130</td>
<td></td>
<td>.0597</td>
</tr>
<tr>
<td>Context Effects</td>
<td>Gy</td>
<td>Gn</td>
<td></td>
<td>By</td>
</tr>
<tr>
<td>Cy</td>
<td>-.0726</td>
<td>.0192</td>
<td></td>
<td>-.0025</td>
</tr>
<tr>
<td>Cn</td>
<td>-.0224</td>
<td>.0756</td>
<td></td>
<td>.1015</td>
</tr>
</tbody>
</table>

### Test order effects:

<table>
<thead>
<tr>
<th></th>
<th>Clinton-Gore</th>
<th></th>
<th>White-Black</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$ (3) = 10.14,</td>
<td>$\chi^2$ (3) = 73.04,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p &lt; .05$</td>
<td>$p &lt; .001$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results: 72 Pew Surveys over 10 years

Quantile - Quantile Test of Repeat Choice Model

$\text{Chi} = 114.37, p = .000$
Assume: One question followed immediately by another with no information in between

\[
\text{Pr}[A \text{ yes and then } B \text{ no}] = p(A_Y B_N) = \left| P_B P_A S \right|^2
\]

\[
\text{Pr}[B \text{ no and then } A \text{ yes}] = p(B_N A_Y) = \left| P_A P_B S \right|^2
\]

**Theorem**: QQ equality

\[
q = \{ p(A_Y B_N) + p(A_N B_Y) \} - \{ p(B_Y A_N) + p(B_N A_Y) \} = 0
\]
<table>
<thead>
<tr>
<th></th>
<th>Gy</th>
<th>Gn</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clinton-Gore</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cy</td>
<td>.4899</td>
<td>.0447</td>
</tr>
<tr>
<td>Cn</td>
<td>.1767</td>
<td>.2886</td>
</tr>
<tr>
<td><strong>Gore-Clinton</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cy</td>
<td>.5625</td>
<td>.0255</td>
</tr>
<tr>
<td>Cn</td>
<td>.1991</td>
<td>.2130</td>
</tr>
<tr>
<td><strong>Context Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cy</td>
<td>-.0726</td>
<td>.0192</td>
</tr>
<tr>
<td>Cn</td>
<td>-.0224</td>
<td>.0756</td>
</tr>
<tr>
<td><strong>White-Black</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wy</td>
<td>.3987</td>
<td>.0174</td>
</tr>
<tr>
<td>Wn</td>
<td>.1612</td>
<td>.4227</td>
</tr>
<tr>
<td><strong>Black-White</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wy</td>
<td>.4012</td>
<td>.1379</td>
</tr>
<tr>
<td>Wn</td>
<td>.0597</td>
<td>.4012</td>
</tr>
</tbody>
</table>

**Test order effects:**
\[
\chi^2 (3) = 10.14, \\
p < .05
\]

**Test QQ equality:**
\[
q = -.003 \\
\chi^2 (1) = .01, p = .91
\]
\[
q = -.02 \\
\chi^2 (1) = .56, p = .46
\]
Results: 72 Pew Surveys over 10 years

Quantile – Quantile Test of QQ Equality

Chi Sq Test
Lack of Fit = 8.73
p = .4625
4. DYNAMICS
COMPARISON OF MARKOV AND QUANTUM THEORIES

Markov

Schrödinger
Markov

\[ N = \text{no. Markov states} \]
\[ p_i = \text{prob state } i \]
\[ \text{Prob}[\text{state } i] = p_i \]
\[ p = [p_i] = N \times 1 \text{ vector} \]
\[ \sum_i p_i = 1 \]

Quantum

\[ N = \text{no. eigen states} \]
\[ \psi_i = \text{amplitude state } i \]
\[ \text{Prob}[\text{state } i] = |\psi_i|^2 \]
\[ \psi = [\psi_i] = N \times 1 \text{ vector} \]
\[ \sum_i |\psi_i|^2 = 1 \]
Markov

\[ T = N \times N \text{ matrix} \]
\[ T_{ij} = \text{prob transit } j \text{ to } i \]
\[ \sum_i T_{ij} = 1 \text{(stochastic)} \]
\[ p(t) = T(t) \cdot p(0) \]

Observe state \( i \) at time \( t \)
(State reduction)
\[ p(t \mid i) = [0,0,\ldots,1,\ldots0]' \]
\[ p(t + s) = T(s) \cdot p(t \mid i) \]

Quantum

\[ U = N \times N \text{ matrix} \]
\[ U_{ij} = \text{amp transit } j \text{ to } i \]
\[ U^*U = I \text{ (unitary)} \]
\[ \psi(t) = U(t) \cdot \psi(0) \]

Observe state \( i \) at time \( t \)
(State Reduction)
\[ \psi(t \mid i) = [0,0,\ldots,1,\ldots0]' \]
\[ \psi(t + s) = U(s) \cdot \psi(t \mid i) \]
Markov

Kolmogorov Eq

\[ \frac{d}{dt} T(t) = K \cdot T(t) \]

Intensity Matrix

\[ K = [k_{ij}] \]

\[ k_{ij} > 0, i \neq j, \]

\[ \sum_j k_{ij} = 0 \]

Quantum

Schrödinger Eq

\[ i \frac{d}{dt} U(t) = H \cdot U(t) \]

Hamiltonian Matrix

\[ H = H^\dagger, \text{ Hermitian} \]
RANDOM WALK MODELS OF DECISION MAKING

Kvam, Pleskac, Busemeyer
(Proceedings of the National Academy of Science)
Random Dot Motion Task

- **T0**: Start of the task
- **T1**: Decision Time
  - 500 ms
- **T2**: Inter-judgment Time
  - 50 / 750 / 1500 ms

**Click (L/R)**
- Respond to the direction of the motion

**Confidence (L/R)**
- Indicate confidence in the decision

*Beep*
- Signal for the next phase of the task
N=7 state Random Walk Model of Confidence
(Toy example, actual model uses N=101 states)
CRITICAL TEST OF MODELS

Condition 1: Measure confidence only at \( t_2 \)
Condition 2: Measure choice at \( t_1 \) and confidence at \( t_2 \)

**Markov**
\[
p(C(t_2) = k \mid \text{Cond2}) = p(C(t_2) = k \mid \text{Cond1})
\]

**Quantum**
\[
p(C(t_2) = k \mid \text{Cond2}) \neq p(C(t_2) = k \mid \text{Cond1})
\]

Statistically test distribution differences using Kolmogorov- Smirnov Statistic
Markov Random Walk

A

Initial State ($t_0$)
Both models start with an initial state centered at 50, indicating uncertainty about the dot motion direction.

B

Quantum Random Walk

C

Dynamics ($t_0 - t_f$)
Dynamics are applied to move the state to its position at $t_f$. Each model uses drift ($\delta$), which moves the state toward the true state of the world, and diffusion ($\sigma^2$), which moves it out over the states indiscriminately.

E

Choice ($t_c$, when prompted)
At $t_c$ in the choice condition, responses are determined by the state. Confidence levels below 50 result in incorrect answers, while those above 50 result in correct ones.

F

G

Dynamics ($t_f - t_c$)
Dynamics are applied again to move the state from its position at $t_f$ to its position at $t_c$. The same drift and diffusion parameters are used. Final distributions for the choice condition are shown.

H

J

Confidence at $t_2$
The state at $t_2$ is determined, and we show the predicted distributions of confidence responses for the choice and no-choice conditions.
Table 1. Summary of model comparison and statistical effects.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Interference*</th>
<th>Second-Stage Processing†</th>
<th>Log Bayes Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.18 [-0.26, -0.11]‡</td>
<td>0.12 [0.08, 0.18] †</td>
<td>212</td>
</tr>
<tr>
<td>2</td>
<td>-0.15 [-0.23, -0.07]‡</td>
<td>0.08 [0.03, 0.14] †</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>-0.15 [-0.22, -0.07]‡</td>
<td>0.01 [-0.04, 0.06]</td>
<td>-131</td>
</tr>
<tr>
<td>4</td>
<td>-0.14 [-0.23, -0.07]‡</td>
<td>0.10 [0.04, 0.15] †</td>
<td>190</td>
</tr>
<tr>
<td>5</td>
<td>-0.11 [-0.19, -0.04]‡</td>
<td>0.07 [0.02, 0.13] †</td>
<td>837</td>
</tr>
<tr>
<td>6</td>
<td>-0.08 [-0.16, -0.01]‡</td>
<td>0.13 [0.07, 0.18] †</td>
<td>223</td>
</tr>
<tr>
<td>7</td>
<td>-0.07 [-0.15, 0.01]</td>
<td>-0.01 [-0.07, 0.05]</td>
<td>-148</td>
</tr>
<tr>
<td>8</td>
<td>-0.05 [-0.14, 0.02]</td>
<td>0.04 [-0.08, 0.10]</td>
<td>339</td>
</tr>
<tr>
<td>9</td>
<td>-0.01 [-0.09, 0.07]</td>
<td>-0.02 [-0.06, 0.04]</td>
<td>150</td>
</tr>
<tr>
<td>Group Level</td>
<td>-0.11 [-0.18, -0.04]‡</td>
<td>0.06 [0.01, 0.12] †</td>
<td>1713</td>
</tr>
</tbody>
</table>

* The mean posterior coefficient and 95% HDI for the main effect of the choice / click manipulation on half-scale confidence.
† The mean posterior coefficient and 95% HDI for the interaction between dot coherence and second stage processing time on full-scale confidence.
‡ 95% highest density interval for the estimate of the corresponding parameter excluded zero.
CONCLUSIONS

• Quantum theory provides an alternative framework for developing probabilistic and dynamic models of decision making

• Provides a coherent account for puzzling violations of law of total probability found in a variety of decision making studies

• Forms a new foundation for understanding widely different phenomena in decision making using a common set of axiomatic principles
“Mathematical models of cognition so often seem like mere formal exercises. Quantum theory is a rare exception. Without sacrificing formal rigor, it captures deep insights about the workings of the mind with elegant simplicity. This book promises to revolutionize the way we think about thinking.”

Steven Sloman
Cognitive, Linguistic, and Psychological Sciences, Brown University

“This book is about why and how formal structures of quantum theory are essential for psychology - a breakthrough resolving long-standing problems and suggesting novel routes for future research, convincingly presented by two main experts in the field.”

Harald Atmanspacher
Department of Theory and Data Analysis, Institut fuer Grenzgebiete der Psychologie und Psychohygiene e.V.

“Much of our understanding of human thinking is based on probabilistic models. This innovative book by Jerome R. Busemeyer and Peter D. Bruza argues that, actually, the underlying mathematical structures from quantum theory provide a much better account of human thinking than traditional models. They introduce the foundations for modeling probabilistic-dynamic systems using two aspects of quantum theory. The first, “contextuality,” is a way to understand interference effects found with inferences and decisions under conditions of uncertainty. The second, “quantum entanglement,” allows cognitive phenomena to be modeled in a non-reductionist way. Employing these principles drawn from quantum theory allows us to view human cognition and decision in a totally new light. Introducing the basic principles in an easy-to-follow way, this book does not assume a physics background or a quantum brain and comes complete with a tutorial and fully worked-out applications in important areas of cognition and decision.

Jerome R. Busemeyer is a Professor in the Department of Psychological and Brain Sciences at Indiana University, Bloomington, USA.

Peter D. Bruza is a Professor in the Faculty of Science and Technology at Queensland University of Technology, Brisbane, Australia.