Advanced tools and concepts for quantum cognition: A tutorial[☆]

James M. Yearsley[a,b]

[a]*Department of Psychology, School of Arts and Social Sciences, City University London, Whiskin Street, London EC1R 0JD, UK*
[b]*JDM Lab, Department of Psychology, Vanderbilt University, Nashville, TN 37240-7817, USA*

Abstract

This tutorial is intended to provide an introduction to some advanced tools and concepts needed to construct more realistic quantum models of decision. The aim is to cover, in a format suitable for researchers with some limited exposure to quantum models of cognition, the ideas of density matrices, POVM type measurements and open system dynamics. The central theme we explore is how we might introduce noise into our quantum models, and the effect this has on model behaviour. These important ideas are likely to be very useful for constructing more realistic cognitive models, but they are generally not covered by introductory accounts of quantum theory. We hope that this tutorial will help to introduce these tools to other researchers interested in constructing quantum models of cognition.

Keywords:

1. Preamble

Recent years have seen a surge of interest in so-called quantum models of cognition and decision making (Busemeyer and Bruza, 2014; Aerts, 2009; Mogiliansky et.al., 2009; Yukalov and Sornette, 2011; Khrennikov, 2010; Pothos and Busemeyer, 2013; Wang et al., 2013). These models are based on the mathematics of quantum probability theory (QT), but abstracted from the usual physical content. These models have arisen in part as a response to the empirical challenges faced by 'rational' decision-making models, such as those based on Bayesian probability theory (such examples are mostly associated with the famous Tversky-Khanaman research tradition. See e.g. Tversky and Kahneman (1974); Chater et al. (2006).) These quantum models posit that, at least in some circumstances, human behaviour does not align well with classical probability theory or expected utility maximisation. However unlike, for example, the fast and frugal heuristics programme (see, e.g. Gigerenzer et al., 2011), quantum cognition aims not to do away with the idea of a formal structure underlying decision-making, but simply to replace the structure of classical probability theory with an alternative theory of probabilities. This new probability theory has features, such as context effects, interference effects and constructive judgments, which align well with psychological intuition about human decision-making. Initial research involving quantum models tended to focus mainly on explaining results previously seen as paradoxical from the point of view of classical probability theory, and there have been a number of successes in this area (Pothos and Busemeyer, 2013; Wang et al., 2013; Trueblood and Busemeyer, 2011; White et al., 2014; Pothos and Busemeyer, 2009; Aerts et al., 2013; Bruza et al., 2015; Blutner at al., 2013; Pothos et al., 2013, 2015). More recently, the focus has switched to some extent to testing new predictions arising from quantum models, and designing better tests of quantum vs classical decision theories (Atmanspacher and Filk, 2010; Yearsley and Pothos, 2014, in press; Wang et al., 2014).

Although good progress has been made in understanding how to build simple cognitive models using QT there has been less work done on developing more sophisticated or realistic models. Toy QT models can provide a good qualitative understanding of some cognitive processes, but if the ultimate goal is to pit QT models against the current best models of cognition and decision making then the sophistication of QT models will need to grow to match such models. This necessitates a move beyond toy models to something more realistic. In addition, many of the current toy QT models have obvious conceptual problems beyond their over-simplicity; one simple example is that typical models for the evolution of cognitive variables are unitary, which means they have no fixed points. Under such a model no amount of evidence presentation can ever lead to a fixed belief state.

The solution to these problems is to move to a slightly more sophisticated framework for QT modelling that can better represent realistic decision making. However there is a challenge here; although this more sophisticated framework is not technically or conceptually more demanding than that employed by toy QT models, it *is* considerably harder to access. Although there are a wealth of introductory accounts of quantum physics that can be easily accessed by cognitive scientists (eg (Isham, 1995; Peres, 1998; Pelnio)) (and even a number specifically written for them, (Busemeyer and Bruza, 2014; Yearsley and Busemeyer, 2015)) it is much harder to find introductory accounts of some of the tools needed to build more realistic QT models.

[☆]Based on a tutorial given at the 37^{th} Annual Cognitive Science Society Meeting, Pasadena, California, USA. July 23^{rd}-25^{th}, 2015.

The purpose of this tutorial is to try and give a brief introduction to some of these tools. The material presented consists mainly of advanced concepts and ideas from the physics literature, some of which have made it into more sophisticated quantum cognitive models, but some of which are yet to find a concrete application. The idea is to present these ideas in a way which makes them accessible to cognitive scientists interested in QT models. A key aim of this tutorial is also to make clear when and why these more sophisticated ideas need to be applied, so that researchers have a better understanding of the limits of the current toy QT models.

Since the potential scope of this tutorial is vast, the presentation style will be somewhat non-standard. Specifically we are aiming for a broad and shallow overview of several different topics, or a sort of 'London Bus Tour' of modern quantum mechanics. We will be satisfied if we can point out some of the major landmarks, and suggest how readers may wish to explore the issues in more depth for themselves, as and when they feel it necessary.

Some more general comments:

1. These notes are not meant as an introduction to the whole of the Quantum Cognition program. In particular we assume prior familiarity with the basics of QT. For those with no previous exposure to QT, we recommend the book by Busemeyer and Bruza (2014) or the recent tutorial Yearsley and Busemeyer (2015) as a good place to start. For more detail about the maths, Isham (1995) is a good reference, alternatively there are many sets of excellent lecture notes available online Pelnio.
2. The mathematics underlying some of the ideas presented here is extremely interesting. However since this is intended as an introductory account we will mostly avoid lengthy proofs or derivations, including them only when we feel they aid understanding.
3. A comment on notation; we will be making use of bra-ket notation throughout. Thus state vectors will be written as $|\psi\rangle$. Also we won't usually write hats on operators or use bold/underline to denote vectors - whether something is an operator, a vector or a number should be obvious from the context. Finally we will adopt the physicists' convention of setting $\hbar$ equal to 1.
4. A note on references; we have tried to include only the most useful references we could find. For the most part this means books or review papers where possible, but we've also included genuine research papers where they are useful/comprehensible. There is pretty much no limit to the number of references one could include, see e.g. Cabello (2000). Two sources are worthy of particular mention; 1) arXiv.org. Cognitive scientists may not be familiar with this, but it's a pre-print archive used by the physics/maths community as a place to upload papers prior to publication. Most are subsequently updated upon publication to reflect the published versions. The upshot of this is that probably the majority of published physics papers, dating back to the late 90's, are available free from this one site. Where we can find it therefore, we have included the arXiv reference alongside the journal info, to make papers easier to track down. 2) The single text used most in putting these notes together is *The Theory of Open Quantum Systems* by H.-P.Breuer and F.Petruccione (2006). This is much more a physics text than a psychology one, but it's nevertheless worth a special mention, since we've consulted it so frequently while preparing this tutorial.

2. Introduction: Noise!

The most commonly encountered toy models of QT are idealisations in a number of ways. The key one for the purpose of this tutorial is that they assume experimenters have perfect knowledge/control over the cognitive state of participants, the form and effect of measurements, and finally the details of any 'evolution' of the state.

In the real word (or even the real lab), things are rarely this simple. We want to show you some tools that can let you generalise the models you've come across so far to apply in more realistic situations. It turns out that doing this will also teach us some profound things about the meaning of the quantum approach to cognition, and how it differs from classical approaches. The theme of this tutorial is therefore 'Noise', specifically 'Noise in the cognitive state', 'Noise in the measurements' and finally 'Noise in the evolution.'

3. Noise in the cognitive state: Density matrices

3.1. Introduction

Let us begin with a simple motivating example. Suppose we wish to perform an experiment in the lab and the expected results depend on whether participants are left or right-handed. Our PhD student collects an equal number of left and right handed participants and lets them into the lab one at a time. Unfortunately the PhD student doesn't tell us which participants are which, so all we know is that there's a 50/50 chance of getting a left/right-handed participant each time. Suppose the cognitive state of the left handed participants is given by $|L\rangle$ and that of the right handed ones by $|R\rangle$ (and that these two states are orthogonal), what is the correct cognitive state to describe our unknown participants?

You might guess the answer is,

$$\psi = \frac{1}{\sqrt{2}}(|L\rangle + |R\rangle) \tag{1}$$

but this turns out not to be correct. You might have guessed this because if I ask "What's the probability that a participant given by this state will say they are left/right-handed if we ask them?" then the answer is;

$$\begin{aligned} p(left) = \langle\psi| P_L |\psi\rangle &= \frac{1}{2}[\langle L| P_L |L\rangle + \langle L| P_L |R\rangle \\ &\quad + \langle R| P_L |L\rangle + \langle R| P_L |R\rangle] \\ &= \frac{1}{2} \end{aligned} \tag{2}$$

and the same for right. (Here $P_L = |L\rangle\langle L|$ etc.)

However this isn't the correct state because what we've made here is a 'quantum' mixture (or superposition) of left and right, whereas what we were really looking for was a classical mixture. $|\psi\rangle$ tells us the participant is in some sense neither left nor right handed [1], at least until we ask, whereas of course what's really happening is that each participant is definitely either left or right handed when they enter the lab, we just don't know which.

In other words, what we want to do is to add in classical uncertainty to our description of a quantum system. This section describes how to do this.

3.2. The Density Matrix ρ

Let's see if we can get a clue about the right answer by looking at the statistics for the outcomes of our experiment on these participants. Suppose our experiment is represented by an operator O, and for left and right handed participants the expected result is l and r respectively. Since we have an equal number of left and right handed participants half the time we will get the result l and half the time we will get the result r. The average outcome across many experiments will therefore be,

$$\begin{aligned}\langle O\rangle &= \frac{1}{2}\langle L|O|L\rangle + \frac{1}{2}\langle R|O|R\rangle, \quad (3)\\ &= \frac{l+r}{2}.\end{aligned}$$

We can write this result in a simpler way by introducing the *density matrix* ρ,

$$\rho = \frac{1}{2}(|L\rangle\langle L| + |R\rangle\langle R|) \quad (4)$$

Then the expected outcome of our experiment can be written as,

$$\langle O\rangle_\rho = \mathrm{Tr}(O\rho) \quad (5)$$

where Tr denotes the *trace* of an operator. The trace of an operator is defined by,

$$\mathrm{Tr}(A) = \sum_i \langle\phi_i|A|\phi_i\rangle \quad (6)$$

where the $\{\phi_i\}$ form an orthonormal basis of the Hilbert space. It is easy to show that if the trace of an operator exists, it is independent of the choice of basis $\{\phi_i\}$[2]. (In terms of matrices, the trace of a matrix is just the sum of the diagonal terms.)

More generally, if we have a classical mixture of possible states $|\psi_\alpha\rangle$ which occur with probabilities ω_α this ensemble can be represented by a density matrix,

$$\rho = \sum_\alpha \omega_\alpha |\psi_\alpha\rangle\langle\psi_\alpha| \quad (7)$$

It turns out that every expression you might have previously encountered in quantum theory has an equivalent in terms of the density matrix. In fact density matrices represent the most general way of writing the equations of quantum theory, and they will prove extremely valuable for the rest of this tutorial. It is therefore worth noting a few properties of the density matrix, and the density matrix analogues of some of the familiar expressions in quantum theory.

Properties of the density matrix:

- It is a Hermitian [3] operator, $\rho^\dagger = \rho$.
- It is normalised in the sense that $\mathrm{Tr}(\rho) = 1$.
- It is a *positive* operator, meaning, $\langle\psi|\rho|\psi\rangle \geq 0, \quad \forall|\psi\rangle \in H.$

These three properties essentially ensure that the eigenvalues of ρ are positive, real numbers which sum to 1, and thus have the interpretation of probabilities.

As we mentioned above, all of the expressions you have encountered so far in quantum theory can be rewritten in terms of the density matrix. For example, from the expression for the time evolution of a vector, $|\psi(t)\rangle = U(t)|\psi_0\rangle$, where[4] $U(t) = e^{-iHt}$ it follows that,

$$\rho(t) = U(t)\rho U^\dagger(t) \quad (8)$$

From this, it is easy to see that the analogue of the Schrödinger equation for a density matrix is [5],

$$\frac{\partial}{\partial t}\rho = -i[H,\rho] \quad (9)$$

This is often known as a *master equation*. Finally if we perform a measurement on the state represented by the density matrix ρ the probability that we will get the answer represented by the projection operator P_a is given by,

$$p(a) = \mathrm{Tr}(P_a\rho) \quad (10)$$

and if we do, the state collapses to the new state,

$$\rho' = \frac{P_a\rho P_a}{\mathrm{Tr}(P_a\rho)}. \quad (11)$$

In the special case of $\rho = |\psi\rangle\langle\psi|$ this is easily seen to be equivalent to the usual expression involving state vectors.

[1] Some people often claim that a state such as $|\psi\rangle$ represents a situation where the participant is both left and right handed at the same time. Similarly, in physics people often say things like "The particle can be in two places at once!" However this isn't really correct. If a system has a property A that means that the state must be an eigenstate of the projection operator P_A onto the subspace associated with that property. Thus if our state represented a participant who was left handed, we would have $P_L|\psi\rangle = |\psi\rangle$. Since this isn't true for P_L or P_R the correct conclusion is that $|\psi\rangle$ represents the state of a participant who is *neither* left nor right handed, rather than one who is somehow both at the same time.

[2] The trace operation has a bunch of fun and useful properties that you can read about in any good text on quantum theory. The key ones for us are firstly that it is cyclic, i.e. $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$ and secondly that for any operator A, $\mathrm{Tr}(A|\psi\rangle\langle\psi|) = \langle\psi|A|\psi\rangle$.

[3] Technically self-adjoint, but the difference isn't important here. Note that the dagger operation means conjugate transpose, i.e. $M^\dagger_{ij} = (M_{ji})^*$.

[4] Assuming a time independent H.

[5] The commutator $[\cdot,\cdot]$, is defined as $[A,B] = AB - BA$

We mentioned above that our original guess at the state for an equal mixture of left and right-handed participants, $\frac{1}{\sqrt{2}}(|L\rangle + |R\rangle)$, wasn't correct. Since this state can also be written as a density matrix, we can compare our guess, ρ_g, with the correct answer, ρ_c. Working in the $\{|L\rangle, |R\rangle\}$ basis, we have;

$$\rho_g = \frac{1}{2}(|L\rangle + |R\rangle)(\langle L| + \langle R|) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \quad (12)$$

$$\rho_c = \frac{1}{2}(|L\rangle \langle L| + |R\rangle \langle R|) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad (13)$$

Comparing the two expressions, we can see that they differ only in their 'off-diagonal' elements. Thus the difference between the classical mixture of left and right handed, and the quantum superposition of left and right handed is in some way encoded in these off-diagonal terms in the density matrix. It is tempting therefore to think that the difference between classical and quantum descriptions of a system can be expressed in this way, and that quantum superpositions can be turned into classical mixtures by somehow removing these terms. We will discuss this further in a later section, but for now note that the situation is a bit more complicated than it seems. For a start, the properties of the density matrix guarantee that it is diagonalisable, i.e. all density matrices are diagonal in some basis. The issue about whether a given density matrix represents a classical or a quantum mixture is therefore more about the basis in which it is diagonal Halliwell (2005); Zurek (1991).

3.3. Using Density Matrices

It might be useful at this point to give a short outline of two ways in which density matrices might be used to construct quantum models of decision. Our motivating example was useful for setting the scene and explaining what density matrices are, but it is obviously unrealistic. What is true however, is that the usefulness of density matrices is primarily in the area of modeling individual differences. We mean this less in the sense of explaining the behaviour of particular participants, and more in the sense of predicting the spread of results, rather than just average behaviour. Most quantum models in the literature are rather simple constructions that are concerned with predicting a particular average behaviour (for example the conjunction fallacy.) An important future direction for research will be understanding the spread of participant behaviours, rather than just the average behaviour. Density matrices allow us to do this in two ways;

- If we happen to know that some individual characteristic is important, and we know the distribution of this characteristic in our testing population, then we can make direct predictions about the average behaviour and the spread of behaviours by encoding these differences as an initial density matrix, in a very similar way to our toy example of left and right handed participants.

- Suppose instead we only think there might be some individual characteristic that is important, but we have no idea about its distribution in our testing population. Well then we can encode the differences in an initial density matrix again, but now leave the distribution of the characteristics as a free parameter, and try to fit this distribution from the data. In other words, if we think different groups of participants might show different behaviours, we can use a density matrix to perform a sort of mixed models analysis, and determine what distribution of individual differences best fits the data.

To the best of our knowledge, neither of these approaches have been explored so far, but they obviously represent important next steps for the QT approach, if our ambition is to produce ever more accurate models.

3.4. The Entropy of a Quantum State

We introduced the density matrix as a way to capture a classical uncertainty about the quantum state. It is therefore natural to ask about the entropy associated with a given density matrix. The entropy of a classical state is a frequently used quantity, and is obviously central to approaches like MaxEnt. Having a quantum analogue is therefore very useful. However before we do this we will first look at a simpler measure of uncertainty, called the 'purity' of a quantum state. This is defined by,

$$\gamma = \mathrm{Tr}(\rho^2) \quad (14)$$

If we write our density matrix in diagonal form, i.e. as,

$$\rho = \sum_i p_i |\phi_i\rangle \langle \phi_i| \quad (15)$$

where the $\{\phi_i\}$ form a complete orthonormal basis, then,

$$\gamma = \mathrm{Tr}(\rho^2) = \sum_i (p_i)^2 \quad (16)$$

Either by diagnalising or by directly squaring the matrix, we can see that,

$$\gamma_g = 1, \quad \gamma_c = \frac{1}{2} \quad (17)$$

where, recall, ρ_g was our guess for the mixture of left and right states, and ρ_c is the correct expression. Density matrices which can be written in the form $\rho = |\psi\rangle \langle \psi|$ always have $\gamma = 1$ and are known as *pure* states, states which cannot be written in this form have $\gamma < 1$ and are called *mixed*. Clearly ρ_g is a pure state, whereas ρ_c is mixed. The purity of a density matrix turns out to be a useful approximate measure of the entropy of the state, but to see this we first need to define the entropy proper.

For a classical probability distribution over a finite set of variables, $\{p_i\}$, the classical Shannon entropy is given by [6],

$$S = -\sum_i p_i \ln(p_i) \quad (18)$$

Now Eq.(15) suggests that we could define the quantum analogue of the Shannon entropy in the same way as Eq.(18), but

[6] We assume here and throughout that $0 \cdot \ln(0) = 0$.

where the p_i are now the 'probabilities' associated with the various basis states $|\phi_i\rangle$. In the basis where ρ is diagonal, this would be equivalent to [7],

$$S = -\mathrm{Tr}(\rho \ln(\rho)) \tag{19}$$

but recall the trace operation is basis independent, thus Eq.(19) is valid generally. It is straightforward to compute the entropies of our two quantum states,

$$S_g = 0, \quad S_c = \ln(2). \tag{20}$$

Eq.(**??**) is know as the von Neumann entropy ((Neilsen and Chuang, 2000)). It is easily seen that in general pure states like ρ_g have zero entropy.

We can now explain briefly one reason why the purity is such a useful measure. Suppose our density matrix is close to being pure i.e. $\rho^2 \approx \rho$. We can Taylor expand the logarithm as,

$$-\ln(\rho) = (1-\rho) + (1-\rho)^2/2 + (1-\rho)^3/3 + ... \tag{21}$$

It follows that

$$S = \mathrm{Tr}(\rho - \rho^2) + \text{ higher order terms} \tag{22}$$

$$= 1 - \gamma + \text{ higher order terms} \tag{23}$$

The quantity $1-\gamma$ is often called the *linear entropy*, as it's the term that comes from the linear expansion of $\ln(\rho)$. The linear entropy is a lower approximation to the von Neumann entropy, but is much easier to calculate, since it doesn't involve diagonalising ρ. In Fig.1 we plot both the von Neumann entropy and the linear entropy as a function of p, for the state,

$$\rho = p\,|L\rangle\langle L| + (1-p)\,|R\rangle\langle R| \tag{24}$$

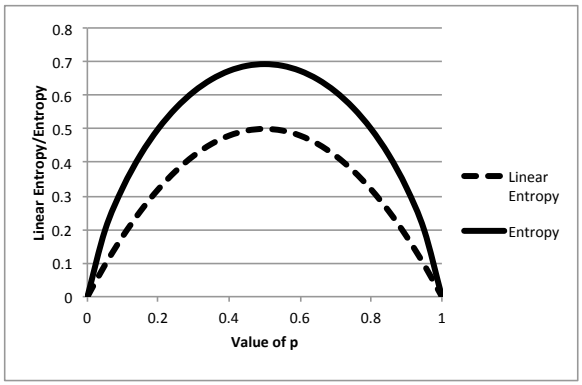

Figure 1: The von Neumann and linear entropies for the state Eq.(24).

For a classical probability distribution, the maximum entropy state is the one with equal probability for any outcome. The quantum analogue of this is a density matrix which is diagonal, and where all the diagonal elements are equal. This state is given (for a Hilbert space of dimension d) by,

$$\rho_{\text{max entropy}} = \frac{1}{d}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \frac{1}{d}\mathbb{1}_d \tag{25}$$

[7] Note that if $A_{ij} = A_i\delta_{ij}$ then $f(A_{ij}) = f(A_i)\delta_{ij}$. i.e. any function of a diagonal matrix is itself a diagonal matrix.

where $\mathbb{1}_d$ is the identity matrix in d dimensions (we will usually drop the subscript d since the number of dimensions should be obvious.) It's easy to see that,

$$S(\rho_{\text{max entropy}}) = \ln(d). \tag{26}$$

which agrees with the classical result for a maximally uncertain state.

3.5. Discussion

The introduction of states with classical as well as quantum uncertainty represents a very significant development in the quantum formalism. We can now go ahead and represent a much more general variety of knowledge states, which probably better reflect the kind of participants we might encounter in a realistic experiment.

However the introduction of these states also gives us a chance to discuss what we think is one of the most powerful a priori reasons for considering quantum models of cognition and decision. At the heart of this argument is the difference between classical and quantum uncertainties. Suppose we have a classical system such as a fair coin, where our best description consists of a probability distribution for the two possible outcomes, head or tails. The probabilistic description reflects the fact that we are uncertain about the outcome of a given coin toss. Classical probability distributions have an associated Shannon entropy, and so classically the uncertainty about the outcome of the coin toss (probabilistic description) is related to a lack of knowledge (entropy) about the state of the system. In other words, classically we have,

Uncertainty about outcomes

$$\Leftrightarrow$$

Lack of knowledge about state

Suppose we want to build a classical cognitive model of a participant's preferences; specifically, let's imagine we want a model of what type biscuit the author will choose to eat from the box he has in front of him, as a reward for finishing this subsection. To simplify matters, suppose there are only two types, milk or dark chocolate. The author professes to be indifferent between them currently, and experience (and the number of each kind left in the box) tells us he is equally likely to choose either variety. A classical description of the author's cognitive state, which aimed to match his behaviour, would have to be a probability distribution which gave equal weight to milk and dark chocolate biscuits.

There's something odd about this though. To be clear, such a description *would* give the correct statistics for choices. However recall our discussion above; this classical probability distribution would have an associated entropy, and should be interpretable in terms of a lack of knowledge about the state. But wait, what information exactly is it that we lack? In the example of a coin toss above, if we toss a coin and ask you to guess heads or tails, there is of course a 'correct' answer, a probabilistic description reflects your lack of knowledge about the true state of the coin. Going back to our cognitive model, the

only *knowledge* you could be lacking here is the author's *true* preference for his next biscuit type. In other words, every time you use a classical probability distribution to describe a system, you assume that there is a *true* state of the system, that you are ignorant of. This might be ok for coins, but it is far from obvious that this makes sense for decision makers. For example, according to a classical theory, the author really do have a definite preference for my next biscuit, he's just not telling himself what it is...

We don't want to build a quantum model of biscuits preference here, but suppose we represent the author's preference as the superposition state $\frac{1}{\sqrt{2}}(|milk\rangle + |dark\rangle)$. This state has the same expected outcomes as the classical probability distribution, but critically it has zero entropy. That is, for a quantum superposition, although the best description of the expected measurement outcomes is probabilistic, there is no extra information that it is possible to gain that would improve your ability to predict a choice. In quantum theory,

Uncertainty about outcomes

$$\Leftrightarrow$$

Lack of knowledge about state.

For the author's money, this is one of the most powerful arguments for using quantum probability theory to model cognition.

Now, time for that biscuit.

3.6. Summary

To recap; we introduced the concept of a density matrix, which can be used to represent a state with both classical and quantum uncertainty. In particular, if we have two different groups of participants, represented by the *pure states* $|P_1\rangle$ and $|P_2\rangle$, then we can form a *mixed state* by taking,

$$\rho = \lambda |P_1\rangle\langle P_1| + (1-\lambda)|P_2\rangle\langle P_2|, \quad 0 \leq \lambda \leq 1. \tag{27}$$

Here λ gives the relative frequencies of the two types of participants in our ensemble. If the two types occur equally frequently, then $\lambda = \frac{1}{2}$.

All of the quantum theory you have encountered before can be rewritten in terms of density matrices, and we gave a few examples of common relations.

We discussed the purity and entropy of a density matrix. Pure states always have zero entropy, but density matrices let us think about techniques such as entropy maximisation in quantum models.

Finally we noted that quantum models break the connection between uncertainty and entropy, and this might represent a powerful argument for their use in cognition.

4. Noise in the measurements: POVMs

4.1. Introduction

Designing good experiments is hard. Errors can creep into experiments because of participant quality, but they can also occur because of experimental design. One type of error occurs in studies where the questions have some sort of time pressure associated with them. Participants are forced to rush somewhat through the questions, and as such they sometimes click the wrong box or make other simple errors. Another type of error might occur when studies are very long, and participants loose focus and start giving inconsistent answers.

Suppose we wish to model an experiment where we have participants express a preference for one of two alternatives, A or B, and that these are exhaustive and exclusive alternatives. In an ideal measurement these would be represented by projection operators $P_A = |A\rangle\langle A|$, $P_B = |B\rangle\langle B|$. Suppose instead our measurement isn't ideal but, intentionally or otherwise, is subject to some noise. This means some participants who really prefer A will select option B, and vice versa.

Let's see how we might model this. What we want is an operator E_A, whose expectation value in the state $|A\rangle$ is close to one, but which also has a non-zero expectation value in the state $|B\rangle$, and likewise for E_B. That is,

$$\begin{aligned} \langle A| E_A |A\rangle &= 1-\epsilon, \quad & \langle B| E_A |B\rangle &= \epsilon, \\ \langle A| E_B |A\rangle &= \epsilon, & \langle B| E_B |B\rangle &= 1-\epsilon. \end{aligned} \tag{28}$$

Where $0 \leq \epsilon \leq 1$ is some small error probability. Let us also assume,

$$\langle A| E_A |B\rangle = 0, \text{ etc.} \tag{29}$$

In the basis $\{|A\rangle, |B\rangle\}$ these operators can therefore be written as,

$$E_A = \begin{pmatrix} 1-\epsilon & 0 \\ 0 & \epsilon \end{pmatrix}, \quad E_B = \begin{pmatrix} \epsilon & 0 \\ 0 & 1-\epsilon \end{pmatrix}. \tag{30}$$

Can we use these operators to describe a measurement process? It is easily seen that they are not projection operators, unless $\epsilon = 0$ or 1, nevertheless they satisfy the following properties,

- They are *positive* operators, which means they have positive eigenvalues.
- They are *complete*, in the sense that $E_A + E_B = 1$.

These properties mean that for any density matrix ρ,

$$0 \leq \mathrm{Tr}(E_A \rho) \leq 1 \tag{31}$$

and

$$\sum_{i=A,B} \mathrm{Tr}(E_i \rho) = 1 \tag{32}$$

The quantities $\mathrm{Tr}(E_i\rho)$ can thus be interpreted as probabilities, and so E_A and E_B are good candidates to describe a measurement process.

But what measurement process do they describe? Well there are many ways to think about this, but probably the easiest is to note that we can write,

$$E_A = (1-\epsilon)P_A + \epsilon P_B, \quad E_B = \epsilon P_A + (1-\epsilon)P_B. \tag{33}$$

In other words, we can write these operators like,

$$E_A = \sum_i p_A(i) P_i \tag{34}$$

where $p_A(i)$ have (loosely) the interpretation of probabilities. So one way to think about these measurements is that instead of performing a measurement P_A, we instead perform one of the possible measurements P_i with some probabilities $p_A(i)$. So these sorts of measurements look like noisy versions of ideal measurements.

4.2. POVMs

E_A and E_B together form a specific example of what are known as *positive operator valued measures* or POVMs for short Busch et al. (1995); Neilsen and Chuang (2000). POVMs are the most general type of measurements that can occur in quantum theory. What we want to do now is present an outline of the general theory of POVM measurements. After this we will go on to discuss some more concrete examples.

The most general description of a measurement process in quantum theory is given in terms of a POVM, which is a set of operators $\{E_i\}$, that satisfy,

- Positivity, $\langle\psi| E_i |\psi\rangle \geq 0, \quad \forall |\psi\rangle \in H$
- Completeness, $\sum_i E_i = 1$.

The probability that a measurement described by E_i gives a positive answer is then given by,

$$p(i) = \mathrm{Tr}(E_i\rho). \tag{35}$$

A given POVM can have many different possible realisations. A realisation ϕ_i is essentially the operation applied to the state $\rho \to \phi_i(\rho)$, so that,

$$\mathrm{Tr}(\phi_i(\rho)) = \mathrm{Tr}(E_i\rho) \tag{36}$$

The simplest realisation of a POVM $\{E_i\}$ probably consists of just taking the operator square roots of the E_i, i.e. writing

$$E_i = M_i^\dagger M_i \tag{37}$$

we have

$$\phi_i(\rho) = M_i\rho M_i^\dagger. \tag{38}$$

The M_i are often called 'measurement operators'. It's easy to see from this why a given realisation of a POVM isn't unique. Suppose we use different measurement operators given by $M'_i = UM_i$ where U is an unitary operator. Then,

$$M_i'^\dagger M'_i = M_i^\dagger U^\dagger U M_i = M_i^\dagger M_i = E_i \tag{39}$$

so these new measurement operators form a realisation of the same POVM, but,

$$\phi'_i(\rho) = UM_i\rho M_i^\dagger U^\dagger = U\phi_i(\rho)U^\dagger \tag{40}$$

so the final state after the measurement is different in the two realisations.

In the rest of these notes we will mostly ignore the issue of multiple realisations, by sticking to the choice $M_i = \sqrt{E_i}$. In practice the appropriate realisation can be determined from the details of the measurement process.

The analogue of the collapse postulate in terms of POVMs is simply that if a measurement of the POVM $\{E_i\}$ yields the outcome i, then the state collapses to,

$$\rho' = \frac{\phi_i(\rho)}{\mathrm{Tr}(\phi_i(\rho))} = \frac{M_i\rho M_i^\dagger.}{\mathrm{Tr}(E_i\rho)}, \tag{41}$$

where the second equality holds for our simple choice of realisation Eq.(37)

To return to our example above, for the POVM E_A, in the basis $\{|A\rangle, |B\rangle\}$ the associated measurement operator will be,

$$M_A = \begin{pmatrix} \sqrt{1-\epsilon} & 0 \\ 0 & \sqrt{\epsilon} \end{pmatrix}, \quad M_B = \begin{pmatrix} \sqrt{\epsilon} & 0 \\ 0 & \sqrt{1-\epsilon} \end{pmatrix}, \tag{42}$$

which is nice and simple.

One feature of POVMs that is worth noting is that they represent measurements that are not perfectly repeatable, in the sense that if we measure a variable and find the value x then immediately measuring the same variable again will not yield the result x with certainty. This is essentially because for the elements of a POVM $E_A^2 \neq E_A$. Suppose we start with an initial state $\rho = diag(1-p, p)$. If we measure the POVM above and get the outcome A then our state collapses to,

$$\rho' = \frac{1}{(1-\epsilon)(1-p)+\epsilon p}\begin{pmatrix} (1-\epsilon)(1-p) & 0 \\ 0 & \epsilon p \end{pmatrix}. \tag{43}$$

if we now perform another measurement of the POVM then the probability we will get the outcome A again is,

$$\mathrm{Tr}(E_A\rho') = \frac{(1-\epsilon)^2(1-p)+\epsilon^2 p}{(1-\epsilon)(1-p)+\epsilon p} \leq 1 \tag{44}$$

for small ϵ and $p < 1$ this goes as,

$$\mathrm{Tr}(E_A\rho') = 1 - \frac{\epsilon}{1-p} + O(\epsilon^2) \tag{45}$$

so that this POVM measurement is not perfectly repeatable.

4.3. Non-Orthogonal Measurements and 'Maybe'

One interesting property of POVMs, as opposed to a description of a measurement process based on projection operators, is that the elements of a POVM need not be orthogonal. This means that we can have more measurement outcomes than there are dimensions in our Hilbert space. One application of this is where we have a two dimensional set of choices, say A or B, but more than two possible responses, say "prefer A", "prefer B" and "don't know". There are doubtless better ways of modelling this situation, but let's follow this through and see what happens. The states associated with each outcome are given by,

$$\begin{aligned} \text{Prefer } A &= |A\rangle \\ \text{Prefer } B &= |B\rangle \\ \text{Don't know} &= \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle) \\ &\quad \text{or } \frac{1}{\sqrt{2}}(|A\rangle - |B\rangle) \end{aligned} \tag{46}$$

They have associated projection operators P_A, P_B, P_+ and P_- in what we hope is an obvious notation. Now these set of projection operators can't form a description of a measurement, because they are not normalised, i.e.

$$\sum_{i=A,B,+,-} P_i = 2 \tag{47}$$

but we can easily turn them into a POVM by normalising. The POVM is therefore given by the set,

$$E_A = \frac{P_A}{2},\ E_B = \frac{P_B}{2},\ E_+ = \frac{P_+}{2},\ E_- = \frac{P_-}{2}. \tag{48}$$

Suppose our state is $|+\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$. Then we can show,

$$p(A) = \frac{1}{4}, \quad p(B) = \frac{1}{4}, \quad p(\text{don't know}) = \frac{1}{2}. \tag{49}$$

It turns out this example is not very realistic (e.g. $p(\text{don't know}) = \frac{1}{2}$ always!) but we hope it shows POVMs have potential for modeling this kind of measurement. We will see a better example in the next section.

4.4. POVMs for Likert scales

Many experimental conditions in psychology involve requesting responses on a Likert scale. Belief and preference strengths are often elicited in this way, but they can be used to measure almost any cognitive variable.

Likert scales pose an interesting problem for quantum models, because it is somewhat unclear how to approach modeling the corresponding variables. Suppose one is interested in a variable x which we are going to take to be a belief that some event will happen, say that it will rain tomorrow. One might imagine measuring participant's beliefs about this event in a number of different ways,

1. Do you think it will rain tomorrow? (Yes, No)
2. Do you think it will rain tomorrow? (Yes, No, Not Sure)
3. On a 1-9 scale, where 1 is certainty of rain, and 9 is certainty of no rain, how likely do you think it is that it will rain tomorrow?
4. What percentage chance do you think there is of rain tomorrow?
5. ...

It is reasonably clear that these are all measuring the same underlying cognitive variable, belief in rain tomorrow. However if one assumes that the different response options are orthogonal then they appear to require vastly different dimensional Hilbert spaces, from 2 to 100 (or even an infinite dimensional space if one allows a continuous response option like a slider.) Our challenge is to come up with a unified way of modeling these different response types as arising from a single quantum variable. There are many way to do this. We will describe one of the simplest[8].

[8] An even simpler way is to ignore the issue and assume participants have direct access to the probability $\text{Tr}(P_x\rho)$. This is frequently done in quantum models. However this is problematic for two reasons; first this probability is an expectation value that only makes sense for an ensemble of systems, which requires that participants have not a single belief state but a whole collection that they can query. In other words, this is not actually quantum theory of a single belief state anymore. Second, because this measurement is not modeled in a proper quantum way, it is unclear what happens if we ask participants to make sequential Likert scale type judgements. What does the state collapse to after the first judgment?

Suppose we allow the different responses on the Likert scale to be non-orthogonal. Suppose we use an N point scale, we can take the responses $|x = 0\rangle = |0\rangle$ and $|x = N\rangle = |1\rangle$ and the intermediate responses to be,

$$|x = n\rangle = \cos\left(\frac{n\pi}{2N}\right)|0\rangle + \sin\left(\frac{n\pi}{2N}\right)|1\rangle \tag{50}$$

These states give rise to a set of projection operators $\{P_n\}$ that can be turned into a POVM by normalising. The probabilities that a measurement of this variable will yield one of the responses 1-9 for the states $|0\rangle$ and $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ are shown in Fig 2, for the choice of a 9 point response scale. Note that there is

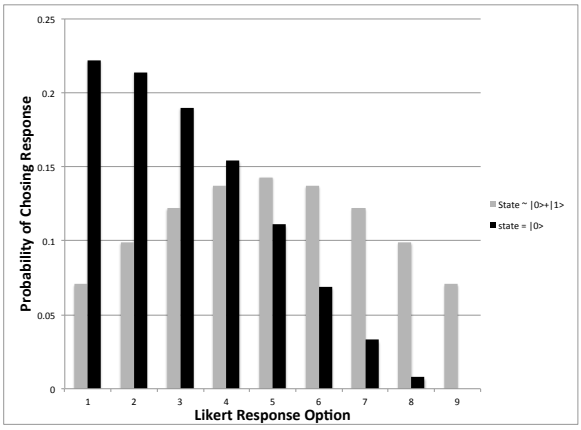

Figure 2: Probabilities for each of the 1-9 options on a Likert scale assuming the POVM based on Eq.(50).

obviously some spread in the choices, however if the true state is $|0\rangle$, say, then the probability of choosing either 1,2 or 3 on the scale is almost 63%.

There are other, perhaps better ways of modeling Likert scale type judgments in quantum theory, in particular this approach rapidly looses its value as the number of responses becomes larger. Nevertheless we hope this semi-realistic example is useful in setting out where POVMs can be used in practice.

4.5. Order effects in noisy measurements

An important question is whether noise in the measurement process spoils the quantum features of that measurement. One example of such a quantum feature is order effects in survey designs (Wang et al., 2014), so we will briefly look at whether noise in the measurements spoils order effects. This is covered in more detail in Yearsley and Busemeyer (2015).

A striking simple example of an order effect is to consider an initial state $|A\rangle$ and two possible projective measurements, P_B

onto the state $|B\rangle$ and P_+ onto the state, $\frac{1}{\sqrt{2}}(|A\rangle + |B\rangle)$. It is easy to see that (working for the rest of this subsection in the basis $\{|A\rangle, |B\rangle\}$),

$$\begin{aligned} p(+ \text{ and then } B) =& \text{Tr}(P_B P_+ \rho P_+) \\ =& \text{Tr}\left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\right) \\ =& \text{Tr}\left(\begin{pmatrix} 0 & 0 \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}\right) = \frac{1}{4} \end{aligned} \tag{51}$$

Whereas,

$$\begin{aligned} p(B \text{ and then } +) =& \text{Tr}(P_+ P_B \rho P_B) \\ =& \text{Tr}\left(\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ =& \text{Tr}\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) = 0. \end{aligned} \tag{52}$$

This striking result arises from the fact that $[P_B, P_+] \neq 0$. So what happens if we replace this set of ideal measurements with a POVM?

We let's replace the projection operators with the following [9],

$$\begin{aligned} P_B \rightarrow E_B &= \begin{pmatrix} \epsilon & 0 \\ 0 & 1-\epsilon \end{pmatrix} \\ P_+ \rightarrow E_+ &= \begin{pmatrix} \frac{1}{2} & \frac{1-2\epsilon}{2} \\ \frac{1-2\epsilon}{2} & \frac{1}{2} \end{pmatrix} \end{aligned} \tag{53}$$

These have associated measurement operators,

$$\begin{aligned} M_B &= \begin{pmatrix} \sqrt{\epsilon} & 0 \\ 0 & \sqrt{1-\epsilon} \end{pmatrix} \\ M_+ &= \begin{pmatrix} \frac{\sqrt{1-\epsilon}+\sqrt{\epsilon}}{2} & \frac{\sqrt{1-\epsilon}-\sqrt{\epsilon}}{2} \\ \frac{\sqrt{1-\epsilon}-\sqrt{\epsilon}}{2} & \frac{\sqrt{1-\epsilon}+\sqrt{\epsilon}}{2} \end{pmatrix} \end{aligned} \tag{54}$$

Now we can see that,

$$\begin{aligned} p_\epsilon(+ \text{ and then } B) =& \text{Tr}(E_B M_+ \rho M_+) \\ =& \frac{1}{4}\left(1 - 2(1-2\epsilon)\sqrt{\epsilon}\sqrt{1-\epsilon}\right) \end{aligned} \tag{55}$$

and

$$\begin{aligned} p_\epsilon(+ \text{ and then } B) =& \text{Tr}(E_+ M_B \rho M_B) \\ =& \frac{\epsilon}{2} \end{aligned} \tag{56}$$

We plot these results against the value of ϵ in Fig.3. The results are interesting. The key is that the difference in the values of the probabilities, ie

$$\delta = p_\epsilon(+ \text{ and then } B) - p_\epsilon(+ \text{ and then } B) \tag{57}$$

[9] Readers are encouraged to convince themselves E_+ is reasonable. Either start with P_B and rotate through $\pi/4$, or consider a combination of P_+ and P_- as in Eq.(34).

(plotted as the dotted line) decreases sharply with increasing ϵ, i.e. with increasing noise. Note however that the value of ϵ is interpretable in terms of the 'error' probability of the measurement. Realistic experiments would probably have values of ϵ in the range 1-5% (as found in Yearsley and Pothos (in press)), and so order effects are still likely to be visible in such experiments, although they might appear smaller than you might have expected.

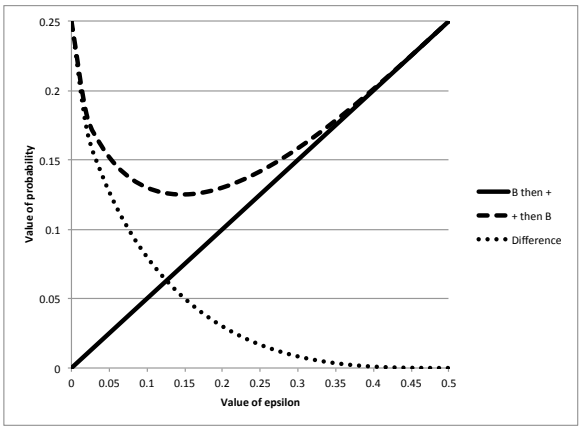

Figure 3: $p_\epsilon(B \text{ and then } +)$, $p_\epsilon(+ \text{ and then } B)$, and their difference, plotted against ϵ.

We don't have space here to pursue this further, but it is clear that small amounts of noise will still allow order effects to be observed, even though very large amounts of noise rapidly kill off such effects. This has important implications for studies looking for these effects in the wild (Wang et al., 2014).

4.6. Summary

We have shown that the description of measurements in quantum theory can be generalised to non-orthogonal sets of measurements. These POVM type measurements can be used to describe noisy realistic measurements, where even participants with a definite knowledge state may not make completely predictable decisions. They can also be used to model situations where there are simply more possible choices than orthogonal states in the space.

We mentioned that one useful way to think about POVMs was as averages over a set of projective measurements, e.g.

$$E_A = \sum_i p_A(i) P_i \tag{58}$$

where the $\{P_i\}$ are a complete and orthogonal set of projection operators and the $p_A(i)$ are positive numbers such that,

$$\sum_A p_A(i) = 1. \tag{59}$$

which ensures the POVMs are normalised.

POVMs are likely to be a very important tool as we strive to make the predictions of the quantum models more accurate. They are also particularly relevant if an experimental set up involves asking participants the same questions repeatedly, see **?** for an example.

5. Noise in the Evolution: CP Maps

5.1. Introduction

The final type of noise we will consider is noise in the evolution of the state. In a typical experiment we manipulate the cognitive state of a participant by presenting some kind of stimuli. Although we might have good control over the stimuli we present, we have much less certainty about how particular participants respond to these stimuli. In addition, we usually assume that different preferences and beliefs are more or less independent of one another, so that, e.g., a model of chocolate biscuit preference can consider this belief state as an isolated system, independent of, e.g. preference for tea or coffee.

In reality, things are not this simple. We therefore need more realistic models of evolution that can help us answer two questions;

1. What effect does the interaction between different beliefs/preferences have on the evolution of a cognitive state?

2. How do we model evolution when we are unsure about the effect of a stimuli on a particular participant?

It will turn out that these two questions have essentially the same answer. In addition an important motivation for considering noisy evolutions turns out to be the following question,

3. How do we model evolutions that are irreversible, or that cause a general initial state to tend towards some final fixed state which is independent of the initial one?

This question arrises because the usual unitary evolution we consider is reversible, ie for any unitary evolution $U(t) = e^{-iHt}$ there is another unitary evolution given by $U^\dagger(t)$ such that $U(t)U^\dagger(t) = U^\dagger(t)U(t) = 1$. There are of course many situations in cognition where we might wish to model evolutions that are (effectively) irreversible, eg decays. This is particularly apparent when we cause a cognitive state to evolve by presenting some stimuli, e.g. some extra information, whose effects we can't undo.

Again, somewhat remarkably, we will find that the answer to question 3 is the same as the answers to questions 1 and 2. We will also find that this way of incorporating noise into cognitive models has some very profound consequences for the 'quantum-ness' of the systems under study.

5.2. CP Maps

Generally speaking, in quantum theory noisy evolutions are motivated by considering a system of interest S which interacts with some other less well controlled system which we call an environment $\mathcal{E}$. We will follow the same reasoning, although it's reasonable to have concerns about how well the physics/cognition analogy works here. In the end though the key point of this subsection is that there is a standard form for these noisy evolutions which guarantees that they make mathematical sense. In practice we just pluck noisy evolutions out of thin air to do a particular job, our only concern being that they match this standard form. However it's useful to have some idea about where they come from [10].

The idea is that we want to study the system S which interacts with an environment $\mathcal{E}$ about which we have little or no interest in or information on. What one does then is to specify the dynamics of the joint system+environment, including information about the evolution and initial states to arrive at a description of the density matrix of the whole, $\rho_{S+\mathcal{E}}$. This density matrix contains information about the environment, which we don't want, so to get at a description of just the system we perform an operation called a *partial trace*, where we sum over the environmental degrees of freedom, essentially throwing away the information we don't want, to leave us with an effective description of the dynamics of the system only in terms of a reduced density matrix ρ_S.

We are interested in the effect this has on the master equation for the system alone, i.e. the evolution equation for the reduced density matrix ρ_S. For the complete density matrix $\rho_{S+\mathcal{E}}$ we have,

$$\frac{\partial}{\partial t}\rho_{S+\mathcal{E}} = -i[H_{S+\mathcal{E}}, \rho_{S+\mathcal{E}}] \tag{60}$$

where $H_{S+\mathcal{E}}$ is the joint Hamiltonian of the system plus environment. When we perform the partial trace to remove the environmental degrees of freedom this becomes,

$$\frac{\partial}{\partial t}\rho_S = -i[H_S, \rho_S] + \mathcal{L}\rho_S \tag{61}$$

where $\mathcal{L}$ is a super-operator which encodes the extra dynamics that come from the system-environment interaction. The most general form this equation can take is the so-called 'Lindblad' form (Lindblad, 1976),

$$\begin{aligned}\frac{\partial}{\partial t}\rho_S =& -i[H_S, \rho_S] \\ &+ \sum_k \left(L_k \rho_S L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho_S - \frac{1}{2} \rho_S L_k^\dagger L_k \right)\end{aligned} \tag{62}$$

where $\{L_k\}$ are a set of operators called the 'Lindblad' operators, which model the effect of the environment.

The key feature of evolution according to the Lindblad equation is that it preserves the properties of the density matrix which are important if it is to describe a real cognitive system. The most important (in the sense of difficult to achieve) property is *positivity*, which recall means that all the eigenvalues of ρ are non-negative. For this reason master equations of the form Eq.(62) are known as 'Completely Positive' or CP-maps [11]. In the next sections we will consider a number particularly useful CP-maps, designed to model specific types of evolution.

5.2.1. Some Extra Detail

We add a few extra points here for interested readers.

[10] It's worth noting that this will be pretty wooly. The full mathematical treatment is complex and irrelevant for usage we want to make of the formalism.

[11] Actually Eq.(62) has an additional property not needed to preserve the properties of the density matrix, which is that it is continuous in time. Evolutions of the form Eq.(62) are therefore only a subset of possible CP-maps.

In the derivation of Eq.(61) we assume that we can separate out the system and environmental degrees of freedom in the system. So we can write the total Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_\mathcal{E}$. We can therefore choose a basis for the Hilbert space which consists of tensor products of basis vectors from the system and environment, i.e. $\left|\phi_{ij}\right\rangle = |S_i\rangle \otimes \left|\mathcal{E}_j\right\rangle$, where $\{|S_i\rangle\}$ form a basis for $\mathcal{H}_S$, and likewise for the environmental degrees of freedom. The partial trace over the environmental degrees of freedom is therefore defined as,

$$\mathrm{Tr}_\mathcal{E}(A) = \sum_j \left\langle \mathcal{E}_j \right| A \left| \mathcal{E}_j \right\rangle \tag{63}$$

and the reduced density matrix for the system alone is given by,

$$\rho_S = \mathrm{Tr}_\mathcal{E}(\rho_{S+\mathcal{E}}). \tag{64}$$

A couple of further points to note. First, most derivations of Eq.(61) for real systems assume that the initial density matrix can be factorised as

$$\rho_{S+\mathcal{E}}(t=0) = \rho_S(t=0) \otimes \rho_\mathcal{E}(t=0). \tag{65}$$

In other words the assumption is that the system and environment were independent to begin with, at time $t = 0$. This can have some funny effects on the dynamics. For example, one of the features of noisy evolutions is that they tend to kill of *entanglement* between different parts of the system. However if you take two systems and allow them to interact, they will tend to become entangled with each other. These models of noisy evolutions can therefore have some funny behaviour at short times, that results from assumptions made about the initial state to simplify the analysis.

Another point to note is that in order to make the transition from Eq.(60) to Eq.(61) tractable one often has to make simplifying assumptions about the dynamics. Typical assumptions are that the interaction between the system and environment is weak, and also that it is Markovian. Thus many explicit models of noisy evolutions are Markovian. If you have good reason to believe the system you are trying to model does not have this property, you should be extra careful in your choice of CP map.

Finally, if you are interested in seeing how a derivation of the master equation for a real system/environment interaction proceeds, I strongly recommend the paper by Halliwell (2007), for a simple introduction. This gives a simplified derivation of the master equation for quantum Brownian motion, that is, the dynamics of a system coupled to a thermal environment. The full blown analysis can be found in the classic paper by Caldeira and Leggett (1983), and also in Breuer and Petruccione (2006); Hu et al. (1992). A nice introductory tutorial can be found in Kryszewski and Czechowska-Kryszk (2008).

5.3. A CP-Map for Irreversible Evolutions

In this section we want to introduce a tractable example of a master equation we could use to describe a real cognitive system. The example we will discuss is a simple two-level system $\{|1\rangle, |2\rangle\}$, that might be used to model a binary variable. For the rest of this section we'll work in this basis.

Now a good model for the noisy evolution of such a system is given by the so-called *Quantum Optical Master Equation* (QOME), which describes the evolution of a two level system interacting with a thermal (i.e. random) system of other systems. The dynamics of this system are described by a master equation of the form Eq.(62) with,

$$L_1 = a\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad L_2 = b\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{66}$$

and with the specific choice $a = \sqrt{\omega N}$, $b = \sqrt{\omega(N+1)}$ where ω and N are constants.

However the full dynamics of the QOME is complex (For a full solution see Yu (1998); Breuer and Petruccione (2006).). So instead of considering the full dynamics of a system interacting with an environment, what we will do here is to use a special case of this master equation to solve an important problem in quantum models of cognition - how do we model an irreversible evolution?

To keep things simple, we will specialise to the following situation: We have an initial state $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$, and we want to imagine evolving the state in an irreversible way, maybe by giving the participants new information that cannot be taken back, so that the state tends towards $|1\rangle$. Unitary evolution won't work here, first since it's reversible, and second since evolving for long enough could cause the state to 'overshoot' and move back towards $|2\rangle$.

It turns out that one solution to this problem is given by a special case of the QOME, with $a = \sqrt{\gamma}$ some constant and $b = 0$, that is,

$$\frac{\partial}{\partial t}\rho = \gamma\left(L\rho L^\dagger - \frac{1}{2}L^\dagger L\rho - \frac{1}{2}\rho L^\dagger L\right) \tag{67}$$

with $L = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, and we've assumed there is no unitary part to the evolution. The solution, for the initial condition given above, is,

$$\rho(t) = \begin{pmatrix} 1 - \frac{1}{2}e^{-\gamma t} & \frac{1}{2}e^{-\frac{\gamma t}{2}} \\ \frac{1}{2}e^{-\frac{\gamma t}{2}} & \frac{1}{2}e^{-\gamma t} \end{pmatrix} \tag{68}$$

As we can easily see, the solution tends to $|1\rangle$ at large times, and it doesn't overshoot. $|1\rangle$ is therefore a *fixed point* of the evolution. This solution therefore describes a state evolving towards $|1\rangle$, and getting there asymptotically. Solutions for decays towards other states can be obtained by applying unitary transformations to the Lindblad operators.

One interesting feature of this evolution is that it does not preserve the purity of the state, and by extension also the entropy. Typically the entropy initially increases, before decreasing again as the state tends to the known final state. The reason for this is subtle but is essentially due to the fact that this system is not closed but is modeled as interacting with a much larger system. The total entropy of system plus environment will always increase, as it should. We plot the behavior of the entropy and purity in Fig.4.

In summary then, this particular example of a master equation can be very fruitfully used to model irreversible evolutions,

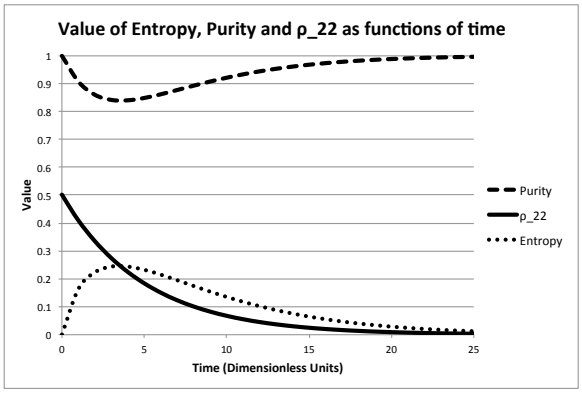

Figure 4: Evolution of the entropy, purity and value of $\rho_{22}(t)$ for the evolution in Eq.(68).

which are common place in the lab, but impossible to model via unitary transformations. There are many other possible master equations for modeling this sort of evolution, but they all have in common that they represent noisy, non-unitary evolution.

5.4. A CP-Map for Uncertain Stimuli: The Montevideo Master Equation

In this section we want to consider a second type of noisy evolution, which is not so readily interpretable in terms of a system/environment interaction. However we shall see that it too can be modeled in terms of a master equation of Lindblad form.

Suppose we wish to model the evolution of a system caused by the presentation of certain stimuli. Suppose we can present our stimuli in a continuous manner, so it could be that the change in the cognitive state depends on the length of time for which the stimuli are presented. Alternatively it could be that the stimuli are discrete, but individually weak, so that presentation of multiple stimuli can be closely approximated by a continuous in time master equation like Eq.(61). Either way, suppose the stimuli are fixed, but their effect on the participants is unknown. We might be able to assume that the effect of the stimuli is on average to produce a shift in a certain direction, but the size of that shift is unknown. This is equivalent to saying that we have a definite evolution but we are unsure, for each participant, how long that participant's state is evolved for.

Specifically, we will assume that the effect of evolution of a state for a time t is not to produce the change $\rho(0) \rightarrow e^{-iHt}\rho(0)e^{iHt}$ but rather,

$$\rho(t) = \int ds p_t(s) e^{-iH(t+s)}\rho(0)e^{iH(t+s)} \tag{69}$$

where $p_t(s)$ is a probability distribution centered around 0, reflecting the distribution of 'evolution times' for our participants, and we have assumed the underlying evolution about which we are uncertain is unitary. If $p_t(s) = \delta(s)$ we recover standard unitary evolution. We have allowed this probability distribution to depend on t also to reflect the fact that the uncertainty in the evolution time, i.e. the width of $p_t(s)$, might depend on the time evolved for, so that longer average evolution times are associated with larger uncertainties.

We want to be able to represent this evolution in the form of a semi-group [12], in other words, if $\rho(t) = \mathcal{L}_t(\rho(0))$, then we want,

$$\mathcal{L}_t(\mathcal{L}_s(\rho(0))) = \mathcal{L}_t(\rho(s)) = \rho(t+s) = \mathcal{L}_{t+s}(\rho(0)). \tag{70}$$

In other words, $\mathcal{L}_t \cdot \mathcal{L}_s = \mathcal{L}_{t+s}$. Writing this in terms of Eq.(69) we see that we can express this evolution in one of two ways,

$$\rho(t_1+t_2) = \int ds p_{t_1+t_2}(s) e^{-iH(t_1+t_2+s)}\rho(0)e^{iH(t_1+t_2+s)} \tag{71}$$

or

$$\begin{aligned}\rho(t_1+t_2) &= \int ds_2 p_{t_2}(s) e^{-iH(t_2+s_2)}\rho(t_1)e^{iH(t_2+s_2)} \\ &= \int ds_1 ds_2 p_{t_2}(s_2)p_{t_1}(s_1) \\ &\quad \times e^{-iH(t_1+t_2+s_1+s_2)}\rho(0)e^{iH(t_1+t_2+s_1+s_2)} \\ &= \int ds dz p_{t_2}(s-z)p_{t_1}(z) \\ &\quad \times e^{-iH(t_1+t_2+s)}\rho(0)e^{iH(t_1+t_2+s)}\end{aligned} \tag{72}$$

Eq.(71) and Eq.(72) are equivalent if,

$$p_{t_1+t_2}(s) = \int dz p_{t_1}(z)p_{t_2}(s-z) \tag{73}$$

which constrains the possible form for $p_t(s)$. One natural choice is the following,

$$p_t(s) = \sqrt{\frac{1}{\pi\sigma t}} e^{-\frac{s^2}{\sigma t}} \tag{74}$$

where $\sigma > 0$ is some constant. This is easily seen to be normalised and to obey Eq.(73). Note also that,

$$\lim_{t\to 0} p_t(s) = \delta(s) \tag{75}$$

in the sense of distributions. ($\delta(s)$ here is the Dirac delta function [13].)

We want to show that this evolution can be written in the form of a master equation. We start with Eq.(69), differentiate both sides with respect to t, and use the very useful property, for small t [14],

$$p_t(s) = \sqrt{\frac{1}{\pi\sigma t}} e^{-\frac{s^2}{\sigma t}} = \delta(s) + \frac{\sigma t}{4}\delta''(s) + ... \tag{76}$$

to obtain,

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= -i[H,\rho] - \frac{\sigma}{4}[H,[H,\rho]] \\ &= -i[H,\rho] + \frac{\sigma}{2}(H\rho H - \frac{1}{2}H^2\rho - \frac{1}{2}\rho H^2)\end{aligned} \tag{77}$$

[12] All physical evolutions have to be expressible in terms of semi-groups, which means that the product of two evolutions is also an evolution. If an evolution also has an inverse, then it is representable as a group, not just a semi-group. Unitary evolutions have this form.

[13] Defined by $\int_{-\infty}^{\infty} dx\delta(x-x_0)\phi(x) = \phi(x_0)$ for any smooth function ϕ.

[14] Despite the fact that this expansion is very useful, we know of very few places where it is discussed. The one reference we have for this is page 703 of Kleinert (2006), but this is a horrific textbook on path integral methods in quantum theory, and hardly a good go to formula book. If in doubt, just integrate this expression against a smooth test function and you can see why the result is true.

So the master equation for this evolution is indeed of Lindblad form, with $L = L^\dagger = \sqrt{\frac{\sigma}{2}}H$.

As an aside, we've called Eq.(77) the 'Montevideo' master equation, because it crops up as part of a particular approach to the foundations of quantum mechanics called the 'Montevideo Interpretation' MVI (2001). The derivation of this equation in that context is rather different, but the net result is much the same.

Let's see what the solution of this master equation looks like, for a particular choice of Hamiltonian H. Let's choose $H = \gamma\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$. This can be thought as generating rotations about the z-axis. We wont derive the solution to the master equation here, although it's pretty easy and we encourage you to try this as an exercise.[15]. The solution is,

$$\rho(t) = \begin{pmatrix}\rho_{11}(0) & e^{-2i\gamma t-\sigma\gamma^2 t}\rho_{12}(0)\\ e^{2i\gamma t-\sigma\gamma^2 t}\rho_{21}(0) & \rho_{22}(0)\end{pmatrix} \tag{78}$$

where $\rho_{11}(0)$ etc are the initial values of those components of ρ.

This solution tells us a lot of interesting information. Firstly, the evolution does nothing to the diagonal components of the density matrix. Secondly the evolution has the effect both of causing an oscillation in the off diagonal components of ρ, and of causing them to decay in magnitude. A nice illustration of this is to choose the state we introduced way back in Section 2, which, recall, was a superposition of left and right handed. Making the identification $|L\rangle \to |A\rangle, |R\rangle \to |B\rangle$ we have,

$$\rho_g = \begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & \frac{1}{2}\end{pmatrix} \tag{79}$$

If we evolve this state according to Eq.(77) we get, for $t \to \infty$,

$$\lim_{t\to\infty}\rho_g(t) = \begin{pmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{2}\end{pmatrix} = \rho_c \tag{80}$$

where, recall, ρ_c was the correct answer for a mixture of left and right handed participants. In other words, this sort of noisy evolution turns quantum superposition states into classically interpretable mixtures!

In summary then, this Montevideo master equation describes a very useful sort of evolution, which corresponds to evolving with a stimuli whose exact effect on participants we do not know. Amongst other things, this has the effect of killing off the off-diagonal terms in the density matrix, and thus making our superposition states look more like classical mixtures.

5.5. Discussion

The result we've just seen, that a particular type of noisy evolution appears to turn quantum superpositions into classical mixtures, turns out to be pretty general. This effect goes by the name of 'decoherence', and is an important part of the story of why we don't see quantum effects like interference in our everyday lives (Zurek, 1991).

In the world of cognitive modeling, this effect is also likely to be significant. Quantum effects are generally not robust when states are allowed to interact with an environment, or when the evolution is otherwise not well controlled. This has two implications; first, it suggests that care might be needed to ensure cognitive states remain quantum during experimental manipulations. Failure to do so could mean no quantum effects are visible. Second, it suggests an explanation for why some cognitive variables do not show quantum effects, perhaps certain preferences/beliefs are just to hard to isolate, and the inevitable interaction between them and other thoughts quickly kills off any quantum behaviour before it can be observed. This is a worthy subject for future research.

5.6. A CP-Map for continuous variables

Most of the variables we study in cognitive science are discrete, either by definition, as for a 'yes', 'no' question, or because having a finite number of response options for a variable such as preference strength is more practical than treating that variable as continuous. However there are some occasions where it might be more appropriate to work with truly continuous cognitive variables. First, it may be the case that we really have a variable that we think of as being continuous, for example preference strength. We could in principle allow a response via a slider or something else with a continuous set of response options, but even if we only allow participants a finite set of response options it may be more appropriate to treat the underlying variable as continuous. After all, it seems reasonable that changing the number of options on a Likert scale should not change the underlying variable being measured. Second, it is hard to construct simple quantum models of discrete variables when the variables can take a large number of values, simply because the Hamiltonian involves many parameters. An alternative is to start with a continuous variable and discretize, this can quickly lead to a sensible Hamiltonian for the discrete variable. A good recent example of this can be found in Kvam et al. (2015).

To cover these cases we need to develop a CP Map for continuous evolutions that is both simple enough to be useful and is reasonably general. Thankfully this problem has already been considered in physics. The general setting is a quantum version of Brownian motion, in a high temperature limit. For full details of this see Caldeira and Leggett (1983); Hu et al. (1992). The key is that under certain conditions this full evolution reduces to,

$$\frac{\partial}{\partial t}\rho = -i[H,\rho] - D[x,[x,\rho]] \tag{81}$$

where H is the Hamiltonian for the closed system, and D is a positive constant that controls the strength of the system-environment coupling. In the (x, y) basis, and assuming a

[15] To solve a matrix differential equation, first write down the differential equations for each component. You will generally then have a coupled set of equations. Clever changes of variables usually allow you to solve these easily.

Hamiltonian of the form $H = \frac{p^2}{2m} + V(x)$, this reads,

$$\frac{\partial}{\partial t}\rho(x,y) = \frac{i}{2m}\left(\frac{\partial^2\rho(x,y)}{\partial x^2} - \frac{\partial^2\rho(x,y)}{\partial y^2}\right) + i(V(x) - V(y))\rho(x,y) - D(x-y)^2\rho(x,y) \quad (82)$$

A good account of a simplified derivation of this equation, suitable for a general audience, can be found in Halliwell (2007). This master equation can be discretized, either to generate a candidate evolution for a discrete variable, or simply to allow for numerical solution. It may then be solved using standard techniques. Analytic solutions also exist for Eq.(82) in some situations.

Eq.(82) may be written in Lindblad form with operators $L \sim x$. Evolution according to this master equation therefore tends to make the density matrix diagonal in the x basis. In other words, if the variable x represents, say, a particular level of confidence as in Kvam et al. (2015), then the effect of this evolution will be to collapse superpositions states involving different levels of confidence.

5.7. Summary

In this section have have explored the idea of noisy evolutions in quantum cognition. The idea behind such evolutions is that interactions between the system under study and other cognitive variables can influence the evolution of the system we are interested in in a profound way. We saw that the most general type of evolutions in quantum cognition are those that can be written in the form of a master equation of Lindblad form. We saw that these evolutions, also called CP-maps, may also be used to describe irreversible changes to the state, and the effect of presenting stimuli of unknown strength.

We also saw that we can use CP-maps to model irreversible evolutions, such as we might realistically come across in the lab. It is an open question whether the sort of models we have presented match data from real experiments, but the ideas seem promising.

Finally we noted that generally noisy evolutions have the effect of diminishing the 'quantum-ness' of a system. This might have implications for attempts to observe quantum effects in more complex cognitive variables. It might also explain why some such variables do not appear to show any quantum effects.

6. Conclusion

That concludes our whistle stop tour of some topics in modern quantum theory! Let's see if we can recap and draw out any major themes.

In the real world, or even the real lab, all cognitive processes involve some level of noise. We've discussed ways in which we can incorporate noise into quantum systems via the state, the measurements and the evolution. Along the way we also learned some practical lessons, such as how to compute the entropy of a state, how to model non-orthogonal measurements, and how to model evolution where we are unsure about the strength of a stimuli.

Overall we hope that the reader has learnt about some interesting tools and ideas that can be applied to modeling realistic cognitive systems using QT. We look forward to seeing these appear in future research.

6.1. Theme 1: QT can be adapted to cover realistic experiments

Realistic experiments involve inhomogeneous groups of participants, experimental error and noise and interactions between the variables we would like to study, and those we are less interested in. QT can be adapted to incorporate all of these things. The way is open therefore to a whole new level of modeling which aims to reproduce not just the average result, but also the statistical distribution of results for a given trial.

Having the tools to model realistic experiments also let's us think about taking QT out of the lab and into the real world. In particular having an idea about how the qualitative behaviour of a quantum system differs when there is noise present should help us decide where QT might be applicable in the wild. The tools we've covered here can also be used to address other problems in quantum cognitive modeling, such as how to model non-orthogonal measurements.

Overall then, this tutorial will hopefully have some practical value for readers who wish to go on to use QT to model decision making. It should also prove useful to those simply reading research papers in quantum cognition in helping to decide whether modeling assumptions are reasonable or not.

6.2. Theme 2: QT features are robust in the presence of small amounts of noise, but not against large amounts

We haven't often mentioned this explicitly, but it is clear that most quantum features are robust against reasonable levels of noise. This is good to know, as it gives us hope that quantum ideas might be applicable in all sorts of real world situations. However it's also very interesting to see how quantum features can break down when there is too much noise, and the description reverts to looking much like a classical one. This has important implications, both on a practical level, for the types of variables we can hope to see QT effects with, and on a more fundamental level, for our understanding of why some variables appear quantum, and some do not.

6.3. Theme 3: Adding noise to QT teaches us interesting things

One thing we hope we hope to have communicated in this tutorial is that exploring the effects of adding noise to a quantum system does more than simply teach us how to model careless participants. In some ways the real structure of QT is only revealed when we introduce density matrices, POVMs and CP-maps. One particularly important thing we learn is about the connection between classical and QT systems. Studying ideal QT systems might lead you to believe that they are diametrically opposed to classical ones, and the only overlap is for idea cases such as sets of commuting operators. However once we learn about adding noise to a QT system, we can appreciate that classical systems are in some general sense a special case of QT ones, and that it is even possible in some sense to dynamically

transition between QT and classical dynamics. Of course, we have only scratched the surface of this subject, and we'd caution against overgeneralising from the simple cases presented here, but nevertheless we hope that through this tutorial you can begin to glimpse the possibility of a unified approach to decision making, that incorporates both the quantum and the classical. Now that would make an exciting research project...

Acknowledgements

JMY was supported by Leverhulme Trust grant RPG-2013-00 and by NSF grant SES-1556415.

This paper evolved from a set of notes written to accompany a tutorial given at the 37^{th} Annual Cognitive Science Society Meeting, Pasadena, California, USA. July 23^{rd}-25^{th}, 2015. I'd like to thank the other organizers of this tutorial, Jennifer Trueblood, Zheng (Joyce) Wang, Peter Kvam and Jerome Busemeyer, for inviting me to contribute.

I'd also like to thank my colleagues at City University, in particular Emmanuel Pothos, Albert Barque Duran and James Hampton, for many useful discussions which have helped the ideas in this tutorial to take shape.

References

Aerts, D (2009). *Quantum structure in cognition.* Journal of Mathematical Psychology, 53(5), 314-348.

Aerts, D., Gabora, L., & Sozzo, S. (2013). *Concepts and Their Dynamics: A Quantum-Theoretic Modeling of Human Thought.* Topics in Cognitive Science, 5(4), 737-772.

Atmanspacher, H, & Filk, T (2010). *A proposed test of temporal non-locality in bistable perception.* Journal of Mathematical Psychology, 54, 314-321.

Atmanspacher, H, & Römer, H (2012). *Order effects in sequential measurements of non-commuting psychological observables*. Journal of Mathematical Psychology, 56(4), 274-280.

Blutner R, Pothos EM, & Bruza P (2013) *A quantum probability perspective on borderline vagueness.* Top Cogn Sci 5(4):711-736.

H.-P.Breuer & F.Petruccione (2006). *The Theory of Open Quantum Systems* (Oxford University Press).

Bruza, P, Kitto, K, Ramm, B. J. & Sitbon, L. (2015) *A probabilistic framework for analysing the compositionality of conceptual combinations*. Journal of Mathematical Psychology, 67, 26-38.

Busch, P, Grabowski, M & Lahti, PJ (1995). *Operational Quantum Physics*. (Springer Berlin Heidelberg).

Busemeyer, JR & Bruza, PD (2014). *Quantum Models of Cognition and Decision.* (Cambridge University Press).

A Cabello, *Bibliographic guide to the foundations of quantum mechanics and quantum information.* Available at arXiv.org quant-ph/0012089.

Caldeira, A & Leggett, A. (1983). *Path integral approach to quantum Brownian motion*. Physica **121A**, 587.

Chater, N, Tenenbaum, JB, & Yuille, A (2006). *Probabilistic models of cognition: Conceptual foundations.* Trends in cognitive sciences, 10(7), 287-291.

Dirac, PAM (1982). *Principles of Quantum Mechanics.* (Oxford University Press).

Gambini,R. Garcia-Pintos, L. P. Pullin, J. (2011) *An axiomatic formulation of the Montevideo interpretation of quantum mechanics*, Studies in the History and Philosophy of Modern Physics, 42, 256-263. If you'd like to know more, a good source of information is www.montevideointerpretation.org. I don't personally subscribe to this approach, but it's interesting.

Gigerenzer, G, Hertwig, R & Pachur, T (Eds.) (2011). *Heuristics: the foundations of adaptive behaviour.* (Oxford University Press).

Halliwell, JJ (2005). *How the Quantum Universe Became Classical.* Contemporary Physics **46**, 93. Also available at arXiv.org quant-ph/0501119.

Halliwell, JJ (2007). *Two derivations of the master equation of quantum Brownian motion*. J.Phys.A **40**, 3067. Also available at arXiv.org quant-ph/0607132.

Hu, BL, Paz, JP & Zang, Y. (1992). *Quantum Brownian motion in a general environment: Exact master equation with nonlocal dissipation and colored noise*. Phys.Rev.D **45**, 2843.

Isham, C (1995). *Lectures on Quantum Theory*. (Imperial College Press.)

Khrennikov A (2010). *Ubiquitous Quantum Structure: From Psychology to Finance* (Springer, Berlin).

Kleinert, H (2006). *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets. 4th Edition* (World Scientific).

Kryszewski, S & Czechowska-Kryszk, J. (2008) *Master equation - tutorial approach*. Available at arXiv.org quant-ph 0801.1757.

Kvam, PD, Pleskac, TJ, YU, S & Busemeyer, JR (2015). *Interference effects of choice on confidence: Quantum characteristics of evidence accumulation.* PNAS, 112, 10645.

Lindblad, G (1976). *On the generators of quantum dynamical semigroups.* Commun.Math.Phys. **48**, 2, 119.

Mogiliansky, AL, Zamir, S, & Zwirn, H (2009). *Type indeterminacy: A model of the KT (Kahneman-Tversky)-man.* Journal of Mathematical Psychology, 53(5), 349-361.

Neilsen, M & Chuang, I (2000). *Quantum Computation and Quantum Information*. (Cambridge University Press.)

Plenio, M. *Quantum Mechanics*. Lecture notes available at www3.imperial.ac.uk/pls/portallive/docs/1/613904.PDF

Peres A (1998). *Quantum Theory: Concepts and Methods*. (Kluwer Academic, Dordrecht, The Netherlands).

Pothos, EM, Barque-Duran, A, Yearsley, JM, Trueblood, JS, Busemeyer, J, & Hampton, JA. (2015). *Progress and current challenges with the quantum similarity model*. Frontiers in Psychology, **6**, 205.

Pothos, E.M. & Busemeyer, J.R. (2009). *A quantum probability explanation for violations of 'rational' decision theory*. Proc. R. Soc. B, 276, 2171.

Pothos EM, Busemeyer JR (2013) *Can quantum probability provide a new direction for cognitive modeling?* Behav Brain Sci 36(3):255-274.

Pothos, EM, Busemeyer, JR & Trueblood, JS (2013). *A quantum geometric model of similarity*. Psych. Rev. 120(3), 679.

Trueblood, J. S. & Busemeyer, J. R. (2011). *A comparison of the belief-adjustment model and the quantum inference model as explanations of order effects in human inference.* Cognitive Science, 35, 1518-1552.

Tversky, A. & Kahneman, D. (1974). *Judgment under uncertainty: Heuristics and biases*. Science, 185, 1124-1131.

Wang, Z & Busemeyer, JR. (2013). *A quantum question order model supported by empirical tests of an a priori and precise prediction.* Topics in Cognitive Science, 5(4), 689-710.

Wang Z, Busemeyer JR, Atmanspacher H, Pothos EM (2013) *The potential of using quantum theory to build models of cognition*. Top Cogn Sci 5(4):672-688.

Wang, Z, Solloway, T, Shiffrin, RM & Busemeyer, JR (2014). *Context effects produced by question orders reveal the quantum nature of human judgments.* PNAS, 111, 9431.

White, L. C., Pothos, E. M., & Busemeyer, J. R. (2014). *Sometimes it does hurt to ask: the constructive role of articulating impressions*. Cognition, 133(1), 48.

Yearsley JM & Busemeyer JR. (2015) *Quantum cognition and decision theories: A tutorial.*, Journal of Mathematical Psychology

Yearsley, JM & Pothos, EM (2014). *Challenging the classical notion of time in cognition: a quantum perspective.* Proceedings of the Royal Society B 381, 1781, 20133056.

Yearsley, JM & Pothos, EM (in press). *Zeno's paradox in decision making*. Proceedings of the Royal Society B. In this paper we ask some participants the same question up to 13 times! (With other things happening in between, of course.) In this set up it matters hugely whether the measurements are *exactly*, or only *approximately* perfect.

Yu, T (1998). *Decoherence and Localization in Quantum Two-Level Systems.* Physica **248A**, 393. Also available at arXiv.org gr-qc/9605071.

Yukalov, VI, & Sornette, D (2011). *Decision theory with prospect interference and entanglement.* Theory and Decision, 70(3), 283-328.

Zurek, WH (1991). *Decoherence and the Transition from Quantum to Classical*. Physics Today **44**, 10, 36. Also available in an updated form online at arXiv.org quant-ph/0306072.