Title: Zeno's paradox in decision making

Authors: James M. Yearsley[a*], Emmanuel M. Pothos[b]

Affiliation:
[a] Department of Psychology, Vanderbilt University, Nashville TN, USA.
[b] Department of Psychology, City University London, EC1V 0HB, UK.

*Correspondence to: james.m.yearsley@vanderbilt.edu

Running head: Zeno in cognition

Abstract
Classical probability theory has been influential in modeling decision processes, despite empirical findings that have been persistently paradoxical from classical perspectives. For such findings, some researchers have been successfully pursuing decision models based on quantum theory. One unique feature of quantum theory is the collapse postulate, which entails that measurements (or in decision making, judgments) reset the state to be consistent with the measured outcome. If there is quantum structure in cognition, then there has to be evidence for the collapse postulate. A striking, a priori prediction, is that opinion change will be slowed down (under idealized conditions frozen) by continuous judgments. In physics, this is the quantum Zeno effect. We demonstrate a quantum Zeno effect in decision making in humans and so provide evidence that advocates the use of quantum principles in decision theory, at least in some cases.

Key Words: Decision making, opinion change, constructive influences, quantum theory.

Introduction

The question of the descriptive and normative foundations of decision making has been a focus of scientific inquiry since antiquity. One influential approach has been classical, Bayesian probability theory. Bayesian principles are supported by powerful justifications (e.g., the Dutch book theorem) and strong, entrenched intuition. Bayesian models are considered normative, that is, they describe how decisions 'should' be taken, given the information available. Although research on rationality typically concerns human decision making, Bayesian principles are often motivated from adaptive considerations, that are equally relevant to human and non-human decision makers (1).

Bayesian cognitive models have been successful (2). However, occasionally, researchers have observed a persistent divergence between Bayesian prescription and behavior. These results are most famously associated with the influential Tversky, Kahneman research tradition; e.g. (3), where the decision makers are humans, but there have also been studies showing other animals, such as macaques, displaying similar violations of Bayesian prescription (4). These findings have created deep theoretical divides, with some researchers rejecting entirely a role for formal probability theory in cognitive modeling.

As long recognized, the Bayesian framework for probabilistic inference is not the only one. We call quantum theory (QT) the rules for assigning probabilities from quantum mechanics, without the physics. QT has characteristics, such as contextuality and interference, which align well with intuition about cognitive processes. Some researchers have been exploring whether QT could provide an alternative, formal basis for cognitive theory (5-10). Note that QT cognitive models are unrelated to the highly controversial quantum brain hypothesis (11). If there is (some) quantum structure in cognition, then cognitive processes must be consistent with the collapse postulate in QT, which requires that the cognitive state changes when a measurement (e.g., decision) is performed to reflect the measurement outcome. The idea that decisions can have a constructive influence is not new (12-13). However, on the assumption of quantum structure in cognition, we are led to the striking prediction that intermediate judgments can inhibit opinion change (in a specific way predicted by QT), even in the presence of accumulating evidence. In physics, it can be predicted that a continuously observed unstable particle never decays (14); this remarkable effect is called the Quantum Zeno (QZ) effect. If a similar effect can be observed in decision making, this would provide compelling evidence for a role for QT in cognitive theory. Note that it has previously been suggested that a version of the QZ effect is present in bistable perception (15), however we aim to improve on this by presenting a formalism more amenable to direct testing.

In our experiments, participants read a story about a hypothetical murder suspect, Smith. Smith was initially considered innocent by most participants. Then, at each time step, participants were presented with an (approximately) identically strong piece of evidence suggesting that Smith was in fact guilty. The task was designed as a generic situation of opinion change, from presented information. We develop a QT model for how the opinion state (regarding Smith's guilt) changes with evidence, and we also construct a Bayesian model of the same process, which matches the QT model in the case of no intermediate judgments. From the QT model, we extract the surprising prediction of a QZ effect when intermediate judgments are made and contrast this with the prediction of the Bayesian model.

The quantum Zeno prediction in decision making

We begin with an idealized model for opinion change in our experiments, designed to illustrate the effect. Consider a 2D quantum system, whose state space is spanned by two orthogonal states *I* and *G,* corresponding to the beliefs that Smith is either Innocent or Guilty. Presentation of evidence is represented by a rotation of the state such that an initial state *I* evolves towards *G*, with time (pieces of evidence).

The probability that a measurement of the state will reveal *I*, at each of $N \geq 1$ judgments at times T/N, $2T/N$ ... T is (assuming a typical time independent Hamiltonian, all derivations in supplementary material):

$$Prob\left(I\ at\ time\,\frac{T}{N} \wedge\ I\ at\,\frac{2T}{N} \wedge \ldots\right) = \cos^{2N}\left(\frac{\gamma}{N}\right). \quad (1)$$

Here γ is a dimensionless constant that encodes the effect of the evidence in the absence of intermediate judgments. As the number of measurements, N, increases, there is a decreasing probability that the system will change from *I* to *G*. As $N \to \infty$, the probability that the system will change state vanishes, even after large times (number of pieces of evidence). This is the famous QZ effect (14), often described informally as proof that 'a watched pot never boils'. (The name comes from the (loose) analogy with Zeno's arrow paradox (16).)

The Quantum Model

The derivation leading to Eq.(1) involves a number of assumptions that will not hold in realistic decision making settings. However we can still predict a weakened QZ effect, as a slowing down (in a specific way) of the evolution of the measured opinion state, even under more realistic conditions. Two assumptions need to be relaxed. First, realistic measurements are not perfectly reliable. For each measurement, there is a small probability that a participant will incorrectly provide a response not matching his/her cognitive state. This is problematic when several identical measurements are made, since error rates may compound. Imperfect measurements require the use of positive-operator valued measures (POVMs), instead of projection operators. Instead of freezing as $N \to \infty$, some evolution may still occur, but it will depend only on details of the imperfect measurements (17).

Second, evolution of cognitive variables will not, in general, be well modeled by a time independent unitary evolution. For the situation of interest, we may still assume the dynamics are approximately unitary (see the supplementary material for more details). However it may be that the weight given by participants to a piece of evidence depends on its position in the sequence of evidence, implying a primacy or recency effect. In order to capture this we must employ time dependent unitary evolution.

A form for the time dependent unitary evolution general enough for our purposes is (15,18)

$$U(t_m, t_n) = \exp(-i\ \sigma_x B(t_m, t_n)),$$

where σ_x is one of the Pauli matrices (19). The function $B(t_m, t_n)$ specifies the angle a participant's cognitive state is rotated through when presented with pieces of evidence t_m through t_n. A form for $B(t_m, t_n)$ involving two parameters is proposed in the supplementary material. If t_m is the time of presentation of the m^{th} piece of evidence, then

$$B(t_m, t_n) = \alpha \sum_{i=m+1}^{n} a_i\ e^{-\beta(i-m-1)^2}$$

Here the a_i represent the strengths of the individual pieces of evidence, as measured in isolation. Thus the first piece of evidence in a sequence is given a weight $\sim a_1$ the second is given weight $\sim a_2 e^{-\beta}$, and so on.

Since we expect the cognitive state to tend towards a fixed point as we accumulate more evidence, it seems natural to assume that presenting a piece of evidence later in a sequence should have a smaller effect on the cognitive state than if the same piece of evidence had been presented earlier. This is functionally equivalent to assuming diminishing returns. However other types of order effect have been observed in studies of belief updating (20), and this form for $B(t_m, t_n)$ can also encode a recency effect, depending on the parameter β.

The effect of imperfect judgments is encoded by a simple POVM operator with one free parameter, ϵ. The parameter ϵ reflects how error-less measurements are. For example, if a participant considers Smith innocent, then the probability of responding innocent is only $1-\epsilon$, leaving a probability to respond guilty of ϵ. Full details are given in the supplementary material.

Using the above, we can show that:

$$Prob(I\ at\ t|I\ at\ 0) = (1-\epsilon)^2 cos^2\big(B(0,t)\big) + \epsilon(1-\epsilon) sin^2\big(B(0,t)\big) \tag{2}$$

Eq(2) allows us to determine ϵ and $B(0,t)$, from empirical classical data on the probability of judging Smith's innocence, assuming innocence initially, and varying the number of pieces of evidence presented (without intermediate judgments).

We can also use Eq(2), together with some assumptions about the way judgments change the cognitive state classically, to construct a Bayesian model of the same decision making process. We will do this below, but we note that in the case of no intermediate judgments the QT and Bayesian models will coincide. This means that we can use data obtained in the absence of any intermediate judgments to fix all the parameters in both the QT and Bayesian models. Our central predictions, of the specific way in which intermediate judgments affect opinion change, will therefore be parameter free.

The Quantum Zeno Prediction

We are now ready to develop the prediction of a QZ effect in this decision making setting. We will show that a participant deciding Smith's innocence will be less likely to change his/her initial opinion as the number of intermediate judgments increases. In the supplementary material we compute the probability of judging innocent at each of the intermediate judgments and the final one (N in total), given an initial innocence judgment. By analogy with the physics case, this can be called survival probability (14). The result is;

$$\begin{aligned} Prob^Q('survival', N) &= Prob\left(I\ at\ \tfrac{T}{N} AND\ I\ at\ \tfrac{2T}{N} AND\ \ldots I\ at\ T\right) = \\ &(1-\epsilon)^{N+1} \prod_{i=0}^{N-1} \cos^2\left(B\left(\frac{iT}{N}, \frac{(i+1)T}{N}\right)\right) \\ &+\epsilon(1-\epsilon)^N\, sin^2\left(B\left(\frac{(N-1)T}{N}, T\right)\right) \prod_{i=0}^{N-2} \cos^2\left(B\left(\frac{iT}{N}, \frac{(i+1)T}{N}\right)\right) + O(\epsilon^2) \end{aligned} \tag{3}$$

The first term in this expression corresponds to the probability that the cognitive state is always consistent with innocent, and all the judgments reflect this. The second term corresponds to possibility that the state changes between the second to last and final judgments, but the participant nevertheless responds 'innocent' due to the imperfect measurements. Further terms would correspond to more judgments not matching the cognitive state, or to the state changing back from innocent to guilty, these terms are negligible compared to those included in Eq.(3). If $\epsilon = 0$, $\beta = 0$, and the a_i's are equal then Eq(3) reduces to Eq(1).

Constructing a matched Bayesian model

The QT model assumes that evidence changes the opinion state (as determined by Eq(2)), that judgments may be imperfect, and that judgments are constructive. The third property is the characteristically quantum one, so with the first two elements, we constructed an alternative, Bayesian model for survival probability. It is helpful to denote by I_B the event where a participant *believes* Smith is innocent, and I_R the event where a participant *responds* that Smith is innocent, and similarly for guilty.

The expression we are interested in is the Bayesian analogue of Eq.(3); the survival probability after T pieces of evidence have been presented, given that N judgments have been made. This is

$$Prob^C('survival', N) = Prob\left(I_R \; at \; time \; T, I_R \; at \; time \frac{(N-1)T}{N}, \ldots I_R \; at \; time \frac{T}{N} \middle| I_R \; at \; 0\right)$$

We want to construct this so that it matches the quantum expression in the case of no intermediate judgments (N=1). We will sketch how to do this here, full details are given in the supplementary materials.

As already noted, because Eq(2) does not involve any intermediate judgments it may be interpreted classically. We can therefore read off,

$$Prob(I_B \; at \; time \; t | I_B \; at \; time \; 0) = \cos^2\big(B(t,0)\big)$$

$$Prob(G_B \; at \; time \; t | I_B \; at \; time \; 0) = \sin^2(B(t,0))$$

$$Prob(I_R \; at \; time \; t | I_B \; at \; time \; t) = (1-\epsilon), \qquad Prob(G_R \; at \; time \; t | I_B \; at \; time \; t) = \epsilon$$

$$Prob(G_R \; at \; time \; t | G_B \; at \; time \; t) = (1-\epsilon), \qquad Prob(I_R \; at \; time \; t | G_B \; at \; time \; t) = \epsilon$$

(since the probabilities for judgments given cognitive states do not depend on the time, we may denote them simply as $Prob(I_R|I_B)$ etc.) The probabilities involving transitions from Guilty cognitive states to Innocent ones are assumed to be 0. We therefore have our Bayesian survival probability for the case of no intermediate judgments.

When there are intermediate judgments made we need to know the appropriate function $B^C(t_m, t_n)$ for the evolution of the state. The form we have been using for $B(t_m, t_n)$ for the QT model is difficult to motivate in the Bayesian case because the strength of the primacy/recency effect depends on the time since the last judgment rather than on the total time, effectively being 'reset' after every judgment. This is very natural from a QT perspective, however the judgments are not expected to have such an effect classically. It is therefore more plausible to consider a slightly different function in the classical case, $B^C(t_m, t_n)$, given by

$$B^C(t_m, t_n) = \alpha \sum_{i=m+1}^{n} a_i \, e^{-\beta(i-1)^2}$$

This differs from $B(t_m, t_n)$ only in the fact that the function multiplying the evidence strength depends only on how many pieces of evidence have been presented before it, and not on whether any intermediate judgments have been made. Note that $B^C(0, t_m) = B(0, t_m)$ since the quantum and classical models should agree in the absence of intermediate judgments. In particular this means fitting either function to the data in the absence of intermediate judgments produces the same set of parameters, α, β for both models.

In fact we could continue to use the function $B(t_m, t_n)$ in the Bayesian analysis if we desire, despite the fact it is poorly motivated. It turns out that the Bayesian model performs better when using $B^C(t_m, t_n)$, so we will work exclusively with this.

We can use the information above to derive a prediction for the Bayesian survival probability. To do so we make two assumptions, first that ϵ is small, and secondly that the probabilities involving transitions from Guilty cognitive states to Innocent ones are negligible. We can then show (details in the supplementary material)

$$
\begin{aligned}
Prob^C('survival', N) &= Prob\left(I_R\ at\ time\ T, I_R\ at\ time\ \frac{(N-1)T}{N}, \ldots I_R\ at\ time\ \frac{T}{N} \middle| I_R\ at\ 0\right) \\
&= (1-\epsilon)^{N+1}\cos^2\left(B^C(0,T)\right) \\
&+\ \epsilon(1-\epsilon)^N \sin^2\left(B^C\left(\frac{(N-1)T}{N}, T\right)\right)\cos^2\left(B^C\left(0, \frac{(N-1)T}{N}\right)\right) + O(\epsilon^2)
\end{aligned}
$$

(4)

The main feature of the Bayesian prediction is a reduction of survival probability with more intermediate judgments, because of a probability of error at each judgment. This contrasts sharply with the QT prediction, Eq.(3). We are now ready to test the Bayesian and QT predictions in a realistic decision making scenario.

We noted above that the Bayesian model does not include constructive influences from intermediate judgments. Would it be possible to include such influences? One way to do this might be to regard the memory of having made a previous judgment of guilt/innocence as additional evidence in favor of that conclusion. At the very least such an approach would be ad hoc, but it would also require fine tuning to ensure such a model reproduced the qualitative features of the QT model. We will not pursue these ideas further here.

<u>Experimental Investigation</u>

Participants

We ran the same experiment twice (Experiment 1 and Experiment 2), with different samples, solely as a replication exercise. Thus, we describe the two experiments together. For Experiment 1, we recruited 450 experimentally naïve participants, from Amazon Turk. Participants were 49% male and 50% female (1% did not respond to the gender question). Most participants' first language was English (98%) and the average age was 34.8. For Experiment 2, we recruited 581 experimentally naïve participants from CrowdFlower. Participants were 39% male and 61% female (<1% did not respond to the gender question). Most participants' first language was English (96%) and the average age was 37.4. Apart from the recruitment process, the experimental materials were identical for both experiments. The experiment lasted approximately 10 minutes; Amazon Turk participants were paid $0.50 and CrowdFlower participants $1.00.

Materials and Procedure

The experiment was implemented in Qualtrics. Participants were first provided with some basic information about the study and a consent form, complying with the guidelines of the ethics committee of the Department of Psychology, City University London. If participants indicated their consent to take part in the study, then they received further instructions (see below), otherwise the experiment terminated.

Our paradigm extends the one of Tetlock (21), which was designed to test for primacy effects in decision making. After the screens regarding ethics information and consent, all participants saw the same initial story, regarding Smith, a hypothetical suspect in a murder: "Mr. Smith has been charged with murder. The victim is Mr. Dixon. Smith and Dixon had shared an apartment for nine months up until the time of Dixon's death. Dixon was found dead in his bed, and there was a bottle of liquor and a half filled glass on his bedside table. The autopsy revealed that Dixon died from an overdose of sleeping pills. The autopsy also revealed that Dixon had taken the pills sometime between midnight and 2 am. The prosecution claims that Smith slipped the pills into the glass Dixon was drinking from, while the defense claim that Dixon deliberately took an overdose."

Participants were then given a short set of questions regarding some details of what they had just read, in order to check that they were engaging with the task. These questions were intended to reinforce memory of the story details and to check for participants who were not concentrating on the experiment. The small number of participants who failed to correctly answer these questions were excluded from subsequent analysis. Participants were then asked whether they thought Smith was likely to be guilty or innocent, based on the information provided in the vignette, and to provide a brief justification for their response, as a further check that they were adequately concentrating on the task and to reinforce memory for the response. After every judgment in the study, participants also saw a screen reminding them of their response. The first response is critical, since all quantum model predictions are based on knowledge of the initial (mental) state. Most participants (Experiment 1: 95%, Experiment 2: 89%) initially assumed innocence, and so we excluded participants who initially assumed guilt. (Those participants in fact saw an analogous experimental procedure, with innocent rather than guilty evidence, however the number of participants involved was too small to allow meaningful conclusions to be drawn.)

Participants were split into six groups. The first group was presented with 12 pieces of evidence suggesting that Smith was guilty (participants were told they would only see evidence presented by the prosecution and not by the defense). Each piece of evidence was designed (and pilot tested) to be individually quite weak (Table S1), but cumulatively the effect was quite strong. In fact, participants were directly told that each piece of evidence would be likely to be weak and/or circumstantial. After reading all 12 pieces of evidence, participants were again asked whether they thought Smith was guilty or innocent, and again asked to justify their choice. Participants in the other five groups were shown the same evidence in the same way, and asked to make the same final judgment, but were also asked to make intermediate judgments (and justify their responses). These intermediate judgments were worded in the same way as the initial and final ones, and were requested at intervals of either 1, 2, 3, 4 or 6 pieces of evidence. A small number of participants gave justifications for their judgments suggesting they were not properly engaging with the task, and were therefore excluded from the analysis.

The order of presentation of the evidence was partly randomized. The pieces of evidence were split into four blocks of three pieces of evidence each. The order of the blocks was fixed, but the order of the pieces of evidence within each block was randomized. The reason we randomized evidence order in this way, rather than say simply randomizing the order of presentation of all pieces of evidence, is that there are a total of 12!, or about 480 million, possible orderings of the evidence, so it is impossible to capture a representative sample of the orderings by simple randomization.

After the main part of the experiment, participants were shown the evidence they had encountered, and were asked to rate the strength of each piece on a (1-9) scale (Table S1).

Results and model fits

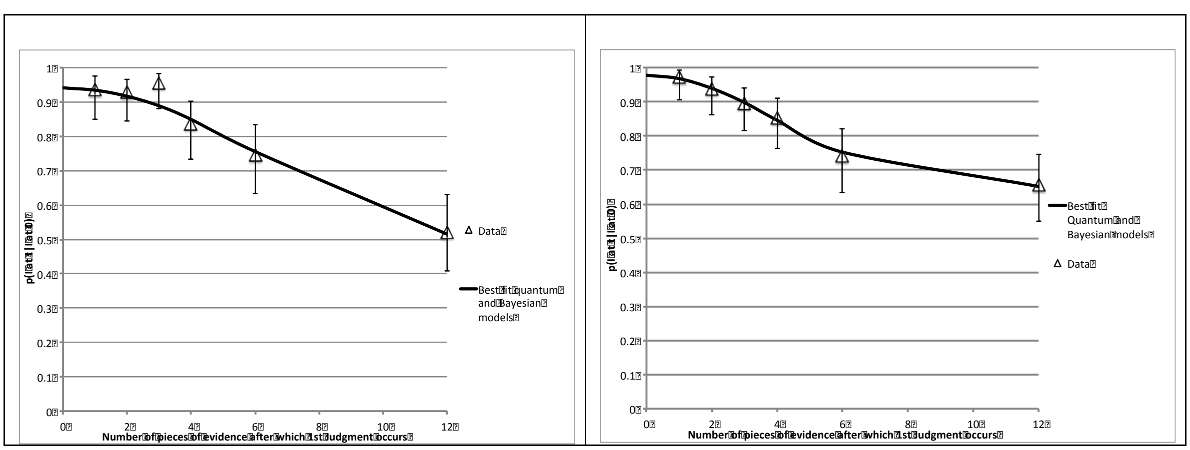
Empirical assessment involved two steps. First, without intermediate judgments (ie at the first judgment made after having seen some evidence) the data is classical and simply informs us how opinion changes with evidence. Using Eq(2), we can determine ϵ and $B(t_m, t_n)$ i.e., the parameter specifying the POVMs for Smith's innocence, guilt and the function specifying the way evidence alters the opinion state (the same parameter values are used in both the Bayesian and QT models). Second, we examined whether the intermediate judgments produce the QZ effect (a slowing down of opinion change, as predicted by the QT model, Eq(3)) or not (in which case the Bayesian model should fit better). The predictions about intermediate judgments from the models were assessed *after* parameter fixing, the first step; they are a priori and parameter free.

In order to determine $B(t_m, t_n)$, we need to know the a_i's for each piece of evidence. These are the parameters indicating the relative strength of each piece of evidence and they were fixed directly, using the participant ratings for each piece of evidence at the end of the task (see supplementary material on fixing the parameters; Table S1). Unfortunately due to an error in the way the experiment was coded, the exact order in which participants saw the pieces of evidence was not recorded. Therefore we set the a_i for each piece of evidence in a given block equal to the average of the reported strengths for the evidence in that block. This is unlikely to cause problems, since the order of presentation of evidence was anyway randomized within blocks.

The best fit parameters were obtained by minimizing the sum of the squared deviations between the predictions of Eq(2) and the data. For Experiment 1, and considering the *t*=3 data point an outlier, the best fit for Eq(2) is obtained with $\alpha = 0.091, \beta = 0.010$ and $\epsilon = 0.030$, giving an R^2 of .996 and a BIC of -27.8 . For Experiment 2, the best fit parameters are $\alpha = 0.114, \beta = 0.0285$ and $\epsilon = 0.0110$, giving an R^2 of 0.99 and a BIC of -23.1. (BICs computed following (22).) The two parameter sets are not equal for the two experiments, a fact we attribute to sampling variation (the demographics of Amazon Turk and CrowdFlower are likely different.) The results of the fitting are shown in Figure 1. (Note that throughout this paper we show error bars corresponding to the 95% Highest Density Interval (HDI) of the posterior distribution for the relevant probabilities, given an initial uniform prior (23).)

For small t, $Prob(I\ at\ t|I\ at\ 0)$ is non-linear and (extrapolated) not equal to 1 at *t*=0. This result justifies our assumption of imperfect measurements. The data from the two experiments show marked differences. In Figure 1a, for large t, $Prob(I\ at\ t|I\ at\ 0)$ is close to linear with increasing *t*. Linearity implies that belief change is proportional to the number of pieces of evidence, which seems an obvious expectation for a rational participant (while the belief state is far from guilty). However, it is unclear whether $Prob(I\ at\ t|I\ at\ 0)$ eventually becomes linear in Figure 1b. Also, more participants gave an initial judgment of 'guilty' in Experiment 2, compared to Experiment 1 (5% vs 11%). Despite distinct behavioral patterns across Experiments 1, 2, Eq(2) provided excellent fits in both cases. Note that the best fit values of β are positive in both cases, confirming our expectation of diminishing returns (equivalently, there is a primacy effect, regarding evidence strength.)

Now that the model parameters have been fixed for both the QT and Bayesian models, we can use Eq(3) and Eq(4) to compute survival probabilities, for different numbers of intermediate judgments.



A.	B.

Figure 1. Setting the parameters (opinion change without intermediate judgments): $Prob(I\ at\ t|I\ at\ 0)$, for the first judgment a participant made, after having seen different numbers of pieces of evidence. A) Experiment 1 (Amazon Turk). Note the obvious outlier at three pieces of evidence. (N=64, 71, 70, 73, 71, 75 for each data point) B) Experiment 2 (CrowdFlower). (N=73, 81, 95, 88, 89, 90). Data points are participant averages and error bars show 95% HDI of the posterior.

Empirical results for $Prob('survival', N)$ clearly favor the QT model (Figure 2). The Bayes factors are 3.4×10^5 for Experiment 1 and 3.2×10^3 for Experiment 2. (Bayes Factors computed following (22).) The classical intuition is reduction of survival probability with more intermediate judgments, because of a probability of error at each judgment. For the QT model, in Experiment 1, we have a clear QZ effect, as survival probability generally increases with N. In Experiment 2, behavior shows a tension between diminishing returns and QZ. With one intermediate judgment, the resetting of diminishing returns means that later pieces of evidence are weighted more strongly than in the case of no intermediate judgments, hence the dip in survival probability. With more intermediate judgments, eventually the QZ effect dominates. The leveling off, or for Experiment 1 the dip in the survival probability for large N is an effect of the imperfect judgments.

There is an alternative test of the QT vs Bayesian models. We can employ Eq(3) and Eq(4) to compute survival probabilities for the condition where there is a judgment after every piece of evidence (number of pieces of evidence presented T, and number of judgments N, vary, but T/N fixed to 1). Again, the data clearly favor the QT model (Figure 3). The Bayes Factors in this case are 8.2×10^9 for Experiment 1 and 1.3×10^9 for Experiment 2.

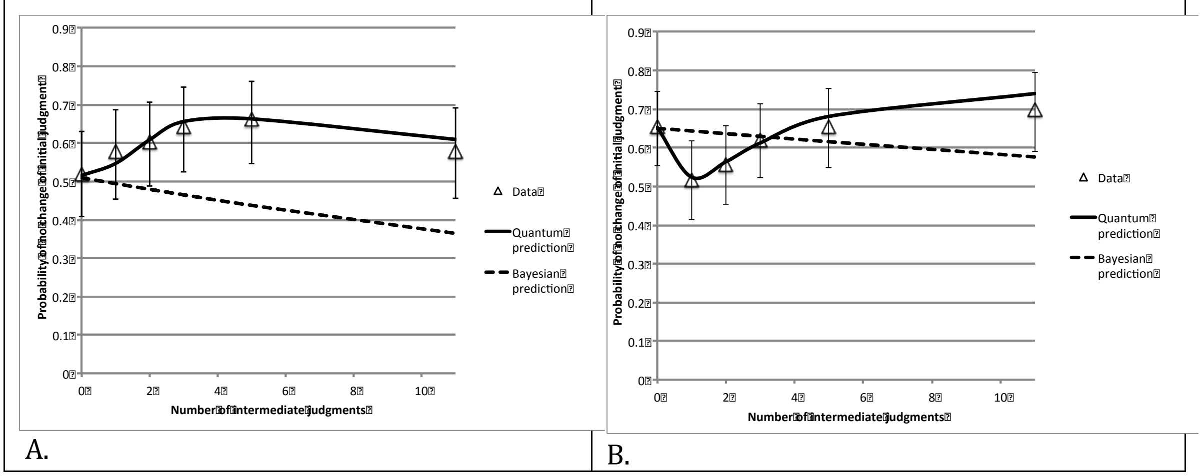


Figure 2. Evaluating the models: Survival probability for N intermediate judgments, for the QT, Bayesisn models, against empirical results (A: Experiment 1, N=75, 71, 73, 70, 71, 64, for each data point; B: Experiment 2, N=90, 89, 88, 95, 81, 73.). Data points are participant averages and error bars show 95% HDI of the posterior.

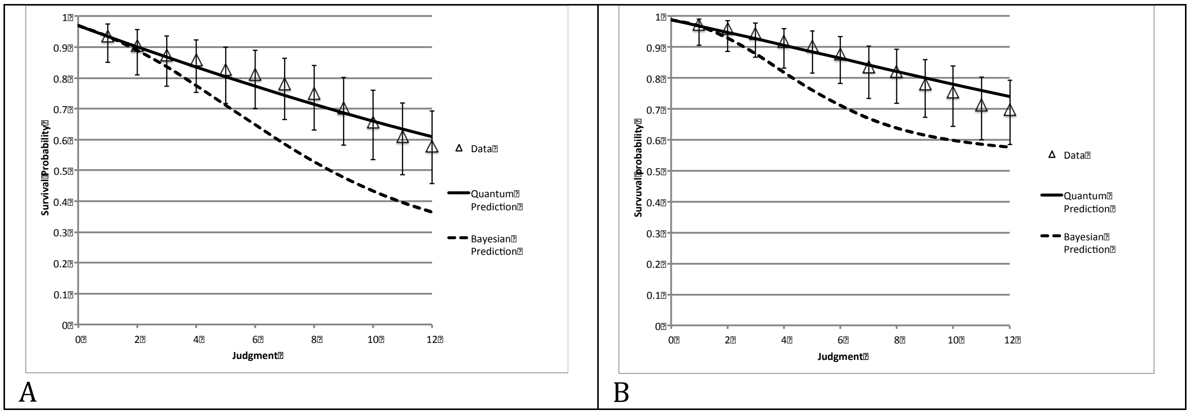


Figure 3. Evaluating the models: Survival probability after each judgment, for the condition with 12 judgments (A: Experiment 1, N=64 for all data points; B: Experiment 2, N=73 for all data points). Data points are participant averages and error bars show 95% HDI of the posterior.

Concluding remarks

Understanding how opinions change (or not) as a result of accumulating evidence is crucial in many situations. We have shown here that opinion change depends not just on the evidence presented, but can also be strongly effected by making intermediate judgments, in the particular way predicted by the quantum model. Because the QT model was fixed with classical data, this striking prediction follows from a structural feature of quantum theory, the collapse postulate, and *not* from parameter fixing. Our results show that decision theory needs to incorporate opinion influences from judgments. They also have practical implications. The employed paradigm has analogies with realistic (e.g., courtroom) assessment of evidence; if e.g. witnesses are expected to reach unbiased conclusions, then the effect of continuous requests for intermediate opinions should be factored in. Likewise, the advent of interactive news web sites (e.g., bbc.co.uk) means that readers can express opinions on news items when reading them, directly and through social media. We raise the possibility that frequent expressions of opinion may prevent change in opinion, even in the presence of compelling contrasting evidence.

More generally, behaviors paradoxical from Bayesian perspectives have often been interpreted as boundaries in the applicability of probabilistic modeling. Strictly speaking this is not true, since one can always augment Bayesian models with extra variables or interactions, however such models may lack predictive power, or simply be too post hoc. The QT cognition program provides an alternative: perhaps some of these paradoxical findings reveal situations where cognition is better understood using QT. Evidence for the collapse postulate in decision making constitutes a general test of the applicability of quantum principles in cognition and adds to the growing body of such demonstrations (8).

While this work has focused on human decision making similar issues apply to animal decision making in general. The adaptive arguments employed to motivate Bayesian principles for humans (1,24) apply equally to non-humans too. Thus, whether Bayesian principles are relevant in animal cognition is an issue of considerable theoretical interest. Is there evidence for constructive influences in animal decision making? A recent study showed that, in the three-door paradigm, pigeons do not show a bias towards repeating a choice when that choice was a guess (25), which is in contrast to behavior seen in humans. This suggests perhaps judgments are less constructive for pigeons than for humans. Clearly the available evidence is far too preliminary to enable strong conclusions. Nevertheless, the demonstration of a QZ effect for humans raises the possibility that a similar effect exists in non-human decision makers. Resolving this question will have potentially ground-breaking implications for understanding the differences between human and non-human mental processes.

Acknowledgments

EMP and JMY were supported by Leverhulme Trust grant RPG-2013-00. Further, JMY was supported by an NSF grant SES-1556415 and EMP was supported by Air Force Office of Scientific Research (AFOSR), Air Force Material Command, USAF, grants FA 8655-13-1-3044. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon. We would like to thank Jerome Busemeyer, David Lusseau and Jenifer Trueblood for useful comments and Thomas Zentall for bringing reference (24) to our attention.

Data Accessibility

Full data sets for the experiments reported in this paper are available via Dryad, doi:10.5061/dryad.n0k69

References and Notes

(1) Anderson, J R (1990). The adaptive character of thought. Hillsdale, NJ: Erlbaum.

(2) Tenenbaum, J.B, Kemp, C., Griffiths, T.L., & Goodman, N. (2011). How to grow a mind: statistics, structure, and abstraction. *Science*, 331, 1279-1285.

(3) Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjuctive fallacy in probability judgment. *Psychological Review*, 90, 293-315.

(4) Hayden, BY, Heilbronner, SR, Nair, AC and Platt, ML (2008). Cognitive influences on risk-seeking by rhesus macaques. *Judgment and Decision Making.* 3, pp. 389-395.

(5) Aerts, D. and Aerts, S. (1995). Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, 1, pp. 85-97.

(6) Busemeyer, J.R. & Bruza, P. (2011). *Quantum Models of Cognition and Decision Making*. CUP: Cambridge, UK.

(7) Haven, E. and Khrennikov, A. (2013). *Quantum Social Science*. CUP: Cambridge, UK.

(8) Pothos, E.M. & Busemeyer, J.R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behavioral & Brain Sciences*, 36, 255-327.

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

(10) Pothos, E. M. & Busemeyer, J. R. (2009). A quantum probability explanation for violations of 'rational' decision theory. *Proceedings of the Royal Society B*, 276, 2171-2178.

(11) Hammeroff, S.R. (1998). Quantum computation in brain microtubules? The Penrose-Hammeroff "orch-or" model of consciousness. Philosophical Transactions of the Royal Society A. 356, 1869-1896.

(12) Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, 25, 638-656.

(13) White, L. C., Pothos, E. M., & Busemeyer, J. R. (2014). Sometimes it does hurt to ask: the constructive role of articulating impressions. *Cognition*, 133, 48-64.

(14) Misra, B. & Sudarshan, E.C.G. (1977). The Zeno's paradox in quantum theory. *Journal of Mathematical Physics*, 18, 756-763.

(15) Atmanspacher, H., Filk, T. & Romer, H. (2004). Quantum Zeno features of bistable perception. Biological Cybernetics, 90, 33-40.

(16) Huggett, N, "Zeno's Paradoxes", The Stanford Encyclopedia of Philosophy (Winter 2010 Edition), Zalta, EN (ed.), plato.stanford.edu/archives/win2010/entries/paradox-zeno/.

(17) Anastopoulos, C. and Savvidou, N. (2006). Time-of-arrival probabilities and quantum measurements. *J. Math. Phys*. 47, 122106.
(18) Yearsley, J.M. & Pothos, E.M. (2014). Challenging the classical notion of time in cognition: a quantum perspective. *Proceedings of the Royal Society B*, 281, 1471-1479.
(19) Isham, C. (1995). Lectures on quantum theory. Imperial College Press. London.
(20) Trueblood, JS and Busemeyer, JR (2013). 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.
(21) Tetlock, P.E. (1983). Accountability and the Perserverence of First Impressions. *Social Psychology Quarterly*, 46, 285-292.
(22) Jarosz, AF and Wiley, J (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *Journal of Problem Solving,* 7, 2-9.
(23) Kruschke, JK and Vanpaemel, W (2015). Bayesian Estimation in Hierarchical Models, in Busemeyer, JR, Wang, Z, Townsend, JT and Eidels, A (Eds) *The Oxford handbook of computational and mathematical psychology*. (OUP, 2015)
(24) Oaksford, M and Chater N (2009). Precis of Bayesian rationality: the probabilistic approach to human reasoning. *Behav. Brain Sci.* 32, 69-120.
(25) Herbranson, WT and Schroeder, J. (2010). Are birds smarter than mathematicians? Pigeons (*Columba livia*) perform optimally on a version of the Monty Hall dilemma. *J Comp Psychol.* 124, 1-13.

<u>Supplementary Material text</u>

The supplementary material text provides detailed description of the quantum and classical models, including derivations of the results in the main text and some additional details regarding the experimental materials.

Derivation leading to Equation (1)

This derivation explains the basic Quantum Zeno effect, under idealized conditions. The idealized situation referred to in the main text concerns a 2D quantum system, evolving under a unitary time independent Hamiltonian.

We prepare our system such that the initial state is $|I\rangle$ at t=0 and let it evolve for a total time $T > 0$. We are interested in the probability that measurements performed on the state at each of the times $T/_N$, $2T/_N$ ...T will confirm that the state is still $|I\rangle$. We have that:

$$Prob\left(I\ at\ time\,\frac{T}{N} \wedge\ I\ at\,\frac{2T}{N} \wedge \ldots\right) = \left|\left(P_I e^{-\frac{iHT}{N}}\right)^N |I\rangle\right|^2 = \left|\left\langle I\left|e^{-\frac{iHT}{N}}\right|I\right\rangle\right|^{2N} \quad \text{(S1)}$$

For a two-level system and a time independent Hamiltonian, transition probabilities typically take the form $\left|\left\langle I\left|e^{-iHt}\right|I\right\rangle\right|^2 = cos^2(E \cdot t)$. In physical applications, E is usually an energy variable. Here, it can be thought of as the average strength of a piece of evidence, since Et is the rotation angle of the mental state, when presented with t pieces of evidence. Eq(S1) then readily leads to the expression, which is Eq(1) in the main text:

$$Prob\left(I\ at\ time\,\frac{T}{N} \wedge\ I\ at\,\frac{2T}{N} \wedge \ldots\right) = cos^{2N}\left(\frac{\gamma}{N}\right),$$

where γ is a dimensionless constant.

Unitary dynamics and POVMs

In this section we motivate the particular choice of dynamics and measurement operators used in the quantum and Bayesian models. We will use this in the next section to derive Eq.(2), which is crucial in the present modeling, since it allows the setting of all parameters with classical data and thus prior to testing for the QZ effect.

In general, in situations such as the one we consider, the most appropriate form of dynamics would be non-unitary. This is because the expected evolution of the mental state is basically like a decay towards a fixed state, the guilty ray, since all the evidence participants encounter is that Smith is guilty and thus, asymptotically, participants must become certain that Smith is guilty.

However, there are two features of our experimental set up that mean that we never need consider mental states close to the guilty ray. First, all participants initially think Smith is innocent, and the evidence we present is designed to be weak, so that the probability that participants judge Smith to be guilty never rises above 50% (as evidenced in the data, e.g., see Figure 2). This means that the evolution by itself never leads to a state close to the guilty state. Thus, the only way a participant's mental state can end up close to the guilty state is by collapsing to this state, if the participant answers that Smith is guilty at one of the intermediate judgments. However, since our analyses were restricted to survival probability, we need not model the further evolution of the mental state after a guilty response. Thus, the only states whose dynamics we are interested in are those far from the guilty state. For these states the fact that the true evolution has a fixed point can, to a good approximation, be ignored, and so the dynamics of such states may be treated as unitary. Of course it is ultimately an empirical question whether this approximation allows for a good fit to the data. In addition, in future work, if it becomes relevant to explore a broader range of experimental manipulations within this paradigm and/or conditions for the mental state, then non-unitary dynamics could be employed.

So far, we have argued that we can model the dynamics of the cognitive state as unitary. However it turns out we need to consider time dependent unitary dynamics in order to capture the expected behavior of the cognitive state. This is essentially because we must allow for the fact that the 'strength' of a piece of evidence may depend on its serial position in the list of evidence presented. It is reasonable (especially in light of earlier remarks about the fact we expect the true evolution to have a fixed point) that we should expect to see a primacy effect, or equivalently diminishing returns, in the weight participants attach to different pieces of evidence. However when we explicitly introduce a form for the evolution in the next section we shall allow for the possibility of either a primacy or a recency effect, and leave it as an empirical question which behavior we see.

We also want to discuss the choice of POVMs to model the measurements. The particular POVMs we use simply model the impact of some noise on the measurements, so that the outcomes are no longer perfectly correlated with the cognitive state. Recall that the projectors representing Innocent and Guilty are given by $P_I = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, P_G = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. The corresponding POVM operators that we use are $E_I = \begin{pmatrix} 1-\epsilon & 0 \\ 0 & \epsilon \end{pmatrix}, E_G = \begin{pmatrix} \epsilon & 0 \\ 0 & 1-\epsilon \end{pmatrix}$, where ϵ encodes the degree of noise. If a participant considers Smith innocent, so that the cognitive state is $|\psi\rangle = \binom{1}{0}$, then the probability of responding innocent is only $1-\epsilon$, leaving a probability to respond guilty of ϵ. Since ϵ is a parameter whose value we estimate from the data it may be that the best fit is provided by $\epsilon = 0$, in which case we recover the usual formalism of projective measurements. Note that the version of the collapse postulate that applies to POVMs is that after a measurement of the POVM E, which yields the answer 'yes', the state changes according to $|\psi\rangle \to \frac{\sqrt{E}|\psi\rangle}{|\sqrt{E}|\psi\rangle|}$. For more on POVMs see (26).

Derivation of Equation (2).

We can now proceed to derive Eq.(2) in the main text. At time 0 participants have not yet heard any evidence and at each time step participants are presented with evidence which supports the possibility of Smith's guilt. The probability that at $t = 0$ a participant initially responds that Smith is innocent is given as:

$$Prob(I\ at\ 0) = \langle\psi|E_I|\psi\rangle = (1-\epsilon)|\langle I|\psi\rangle|^2 + \epsilon|\langle G|\psi\rangle|^2 \qquad \text{(S2)}$$

where E_I is the POVM for innocent. This expression tells us that any participant who answers innocent for this initial judgment (before encountering any evidence) may be assumed to be in state $|I\rangle$ with probability $1-\epsilon$ and in state $|G\rangle$ with probability ϵ.

The general form of the transition probability for a time-dependent Hamiltonian is given by $Prob(I\ at\ time\ t) = \left|\left\langle I\middle|e^{-i\int_0^t dsH(s)}\middle|\psi\right\rangle\right|^2$. Then, the probability that a participant answers innocent after seeing t pieces of evidence, without any intermediate judgments, given an initial response of innocent, is

$$Prob(I\ at\ t|I\ at\ 0) = \frac{\left|\sqrt{E_I}e^{-i\int_0^t ds\,H(s)}\sqrt{E_I}|\psi\rangle\right|^2}{\left|\sqrt{E_I}|\psi\rangle\right|^2} \approx (1-\epsilon)\left|\sqrt{E_I}e^{-i\int_0^t ds\,H(s)}|I\rangle\right|^2$$

(S3)

To progress, we must make some assumptions regarding the Hamiltonian, $H(t)$. The Hamiltonian for any system in a two-dimensional Hilbert space can be written as a sum of the identity operator plus the three Pauli matrices, each with a time-dependent prefactor. As argued elsewhere (15, 18), it is reasonable to simplify the general expression for the time-dependent Hamiltonian of cognitive bivalued systems to $H(t) = b(t)\sigma_x = b(t)\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$, where $b(t)$ is a function of time. Let us next define $B(t_m, t_n) = \int_{t_m}^{t_n} ds\, b(s)$, which incidentally is dimensionless. Then, Eq(S3) can be written as

$$\begin{aligned} Prob(I\ at\ t|I\ at\ 0) &= (1-\epsilon)\left|\sqrt{E_I}e^{-i\int_0^t ds\,H(s)}|1\rangle\right|^2 \\ &= (1-\epsilon)\left|\sqrt{E_I}\left(\mathrm{I}\cdot cos\big(B(0,t)\big) - i\sigma_x\, sin\big(B(0,t)\big)\right)|1\rangle\right|^2 \\ &= (1-\epsilon)^2 cos^2\big(B(0,t)\big) + \epsilon(1-\epsilon)sin^2\big(B(0,t)\big) \end{aligned}$$

which is Eq(2) in the main text.

Understanding the function $B(t_m, t_n)$, fixing it from data, and the Interpretation of the parameters

Both the quantum and classical models for opinion change involve the parameter ϵ, which takes into account erroneous responses, and the function $B(t_m, t_m)$, which tells us how the opinion state changes with accumulating evidence. In this section we describe how the function $B(t_m, n)$ can be specified, how to estimate it from empirical data, and how to interpret its parameters.

Recall, the function $B(t_m, t_n)$ controls the change of the mental state, as a result of considering $t_n - t_m$ pieces of evidence, assuming that a judgment was made at t_m. Therefore a naïve guess at this function would simply be the sum of the relative strengths of all pieces of evidence considered, multiplied by an overall constant, i.e.

$$B(t_m, t_n) =?\, \alpha \sum_{i=m+1}^{n} a_i$$

However the weight given to a piece of evidence may depend on its position in the sequence. Pieces of evidence that come later after a judgment may have less impact on the opinion state than pieces of evidence that come immediately after a judgment, or vice versa. Thus a better choice is,

$$B(t_m, t_n) = \alpha \sum_{i=m+1}^{n} a_i \, g(t_i - t_{m+1})$$

where the function $g(t)$ is a monotonic function of t. The choice of argument is made so that $B(0, t_1) = \alpha \, a_1 g(0)$, and we take $g(0) = 1$ by convention.

Note that the argument of $g(t)$ reflects the number of pieces of evidence seen since the last judgment was made, not the total number of pieces of evidence seen. This is very natural in the quantum model, since the idea is that the process of making a judgment 'collapses' the knowledge state back to the initial state (assuming an 'innocent' judgment.) This implies the state post-judgment should have the same sensitivity to evidence as the initial state, and so any primacy/recency effects should be reset. However this argument cannot be made in a Bayesian model, since 'collapse' is a characteristically quantum feature. Therefore the Bayesian model will involve a slightly different function, $B^C(t_m, t_n)$, where

$$B^C(t_m, t_n) = \alpha \sum_{i=m+1}^{n} a_i \, g(t_i - t_1)$$

There are many choices for the function $g(t)$. We will make the choice $g(t_i - t_{m+1}) = e^{-\beta(i-m-1)^2}$, so that overall we have:

$$B(t_m, t_n) = \sum_{i=m+1}^{n} \alpha \, a_i \; e^{-\beta(i-m-1)^2}$$

$$B^C(t_m, t_n) = \sum_{i=m+1}^{n} \alpha \, a_i \; e^{-\beta(i-1)^2} \qquad \text{(S4)}$$

A positive value of β corresponds to a primacy effect, or diminishing returns, whereas a negative value of β corresponds to a recency effect. This form for $g(t)$ may be motivated by considering a continuous analogue of the process of evidence presentation. Thus, our choice of $B(t_m, t_n)$, involves two free parameters, α, β. Note that there is no fitting regarding the relative strength parameters in Eq(S4), a_i. For a particular piece of evidence i, $a_i = \frac{average\ strength\ for\ evidence\ i}{average\ strength\ of\ all\ pieces\ of\ evidence}$, where both averages are across participants. Crucially the fact that we have reduced the determination of the functions $B(t_m, t_n)$ and $B^C(t_m, t_n)$ to the identification of two parameters means we can fix $B(t_m, t_n)$ and $B^C(t_m, t_n)$given data on $B(0, T)$, which in turn means we can fix it from data which does not concern intermediate judgments. The relative strength of the pieces of evidence, ie the a_i are given in Table 1S.

The parameter α is simply a factor that converts between evidence strength and angle of rotation of the opinion state. It is related to the overall strength of the prosecution's case, but it does not have a particularly interesting interpretation.

The parameter β is more interesting. Its inverse square root indicates the number of pieces of evidence after which the primacy or recency effect starts to have a large impact on the effect of additional evidence. For example, in Experiment 1, the best fit was for $\beta = 0.01$. This tells us that diminishing returns starts to play a role after around 10 pieces of evidence, so we would not expect to see much impact from this in the results. This is evident in Figure 2A, where we see a pure QZ effect. In contrast, in Experiment 2 the best fit was for $\beta = 0.0285$. This suggests diminishing returns should start to have an impact on behavior, after about 6 pieces of evidence. We can see this both in Figure 1B, where there is an obvious change in behavior from 6 to 12 pieces of evidence, and also in Figure 2B. In Figure 2B the noticeable dip in survival probability takes place between one judgment (i.e., only one judgment after all evidence has been presented) and two judgments. This is equivalent to considering the evidence either as one group of 12 pieces (evidence after 6 pieces would have a low impact, broadly speaking) or as two groups of 6 pieces of evidence (according to the quantum model, in this case, after 6 pieces of evidence and one judgment, the following 6 pieces of evidence would also be taken into account in the same way as the original 6; hence, the survival probability drops – more bias that Smith is guilty).

The best fit value for ϵ was approximately 3% in Experiment 1 and 1% in Experiment 2. This means that a participant whose cognitive state is perfectly aligned with the innocent ray may still have a $\approx$ 1% or 3% chance of answering that Smith is guilty, when queried. While this does not appear high for any individual judgment, in an experiment which employs more than two or three judgments, the cumulative error rate can quickly increase beyond 5%. Therefore, with multiple judgments, even in the presence of a simple procedure and very clear instructions (as in the present work), the possibility that participants respond incorrectly (i.e., in a way inconsistent with their mental state) needs to be incorporated in any modeling. The difference in the value of ϵ between Experiment 1 and Experiment 2 explains why there is a dip in survival probability for large N in Experiment 1 (Figure 2A) but this is not observed in Experiment 2 (Figure 2B).

Computing the (quantum) survival probability, for N intermediate measurements (Equation 3)

This section presents the derivation for the quantum survival probability. Following the usual convention in this work of denoting innocence with $|I\rangle$, we have that:

$$Prob('survival', N) = Prob\left(I\ at\ \frac{T}{N} AND\ I\ at\ \frac{2T}{N} AND\ \ldots I\ at\ T\right) \approx$$

$$(1-\epsilon)\left|\prod_{j=0}^{N-1}\sqrt{E_I}\exp\left(-iB\left(\frac{jT}{N},\frac{(j+1)T}{N}\right)\sigma_x\right)|I\rangle\right|^2 =$$

$$(1-\epsilon)\left|\prod_{j=0}^{N-1}\sqrt{E_I}\left(\mathrm{I}.\cos\left(B\left(\frac{jT}{N},\frac{(j+1)T}{N}\right)\right) - i\sigma_x.\,sin\left(B\left(\frac{jT}{N},\frac{(j+1)T}{N}\right)\right)\right)|I\rangle\right|^2$$

(S5)

These probabilities are quite complicated and it is not necessary to give the full expression for every value of *N* here. However, we can simplify them quite considerably by noting that both ϵ and $sin(B(t_i, t_j))$ are small compared to 1. Doing this allows us to write (this is Eq(3) in the main text):

$$\begin{aligned} Prob('survival', N) &= Prob\left(I\ at\ \frac{T}{N} AND\ I\ at\ \frac{2T}{N} AND\ \ldots I\ at\ T\right) \\ &= (1-\epsilon)^{N+1}\prod_{i=0}^{N-1}\cos^2\left(B\left(\frac{iT}{N},\frac{(i+1)T}{N}\right)\right) \\ &+ \epsilon(1-\epsilon)^N sin^2\left(B\left(\frac{(N-1)T}{N},T\right)\right)\prod_{i=0}^{N-2}\cos^2\left(B\left(\frac{iT}{N},\frac{(i+1)T}{N}\right)\right) + O(\epsilon^2) \\ &+ O(sin^4) \end{aligned}$$

(3)

Note that Eq(3) has a reasonably clear interpretation. The first term is the probability that the state never changes, multiplied by the probability that the *N* imperfect measurements all come out in the expected way (i.e., that Smith is innocent). The second term represents the probability that the state changes between the second to last and last measurements, but that the last measurement fails to detect this change. Further terms either represent earlier changes in the state, and so more failed detections, or the state changing back to innocent from guilty (the probability for this last possibility is expected to be negligible for other reasons, since a participant who thinks Smith is guilty is very unlikely to revert and respond that Smith is innocent, after seeing more guilty evidence).

Bayesian survival probability

To derive a Bayesian expression for survival probability, we will assume that the process of making a judgment does not affect the mental state, but, as judgments are imperfect, there is a small probability, ϵ, of making incorrect responses (that is, providing an answer which does not reflect the mental state).

As noted in the main text, much of the information we need to build a Bayesian model can be extracted from Eq(2). Recall that we denote by I_B the event where a participant *believes* Smith is innocent, and I_R the event where a participant *responds* that Smith is innocent, and similarly for guilty. Then from Eq(2) we have,

$$Prob(I_B\ at\ time\ t|I_B\ at\ time\ 0) = \cos^2\big(B(0,t)\big)$$
$$Prob(G_B\ at\ time\ t|I_B\ at\ time\ 0) = \sin^2(B(0,t))$$
$$Prob(I_R|I_B) = (1-\epsilon), \qquad Prob(G_R\,|I_B\,) = \epsilon$$
$$Prob(G_R|G_B) = (1-\epsilon), \qquad Prob(I_R|G_B) = \epsilon$$

The probabilities involving transitions from Guilty cognitive states to Innocent ones are assumed to be 0, as in the quantum model.

The Bayesian survival probability is equal to,

$$Prob^C('survival', N) = prob\left(I_R\ at\ time\ T, I_R\ at\ time\ \frac{(N-1)T}{N}, \dots I_R\ at\ time\ \frac{T}{N}\middle|I_R\ at\ 0\right)$$

We need two assumptions to allow us to write this in terms of quantities we know. The first is that ϵ is small, and the second is that transition probabilities from G_B to I_B are small. The first of these is justified by appeal to the data, the second by the nature of the empirical set up, since we only present evidence implying Smith's guilt. Given these two assumptions, we can show,

$$Prob\left(I_R\ at\ time\ T, I_R\ at\ time\ \frac{(N-1)T}{N}, \dots I_R\ at\ time\ \frac{T}{N}\middle|I_R\ at\ 0\right)$$
$$\approx (1-\epsilon)^{N+1}\ Prob\left(I_B\ at\ time\ T, I_B\ at\ time\ \frac{(N-1)T}{N}, \dots I_B\ at\ time\ \frac{T}{N}\middle|I_B\ at\ 0\right)$$
$$+\,\epsilon(1-\epsilon)^N\ Prob\left(G_B\ at\ time\ T, I_B\ at\ time\ \frac{(N-1)T}{N}, \dots I_B\ at\ time\ \frac{T}{N}\middle|I_B\ at\ 0\right)$$

This follows because the probability of transitioning back to I_B from G_B is essentially 0, and it is very unlikely that the state G_B is incorrectly classified by more than one judgment. Thus the only non-negligible possibility other than that the cognitive state was always aligned with innocent is that the state changed between the penultimate and final judgments.

Next, it is easy to see that,

$$Prob(\dots, I_B\ at\ time\ t_i, I_B\ at\ time\ t_{i-1}, \dots I_B\ at\ time\ t_1|I_B\ at\ 0)$$
$$\approx Prob(\dots, I_B\ at\ time\ t_i|I_B\ at\ 0),$$

which follows because we are assuming the transition probabilities from G_B to I_B are small, so that if the state is I_B now, it is very unlikely to have been G_B at any time in the past. The survival probability then reduces to,

$$Prob^C('survival', N) = \approx (1-\epsilon)^{N+1}\, Prob(I_B \; at\; time\; T | I_B \; at\; 0) + \epsilon(1-\epsilon)^N\, Prob\left(G_B \; at\; time\; T, I_B \; at\; time \frac{(N-1)T}{N} \middle| I_B \; at\; 0\right)$$

We can also write,

$$Prob\left(G_B \; at\; time\; T, I_B \; at\; time \frac{(N-1)T}{N} \middle| I_B \; at\; 0\right) = Prob\left(G_B \; at\; time\; T \middle| I_B \; at\; \frac{(N-1)T}{N}\right) Prob\left(I_B \; at\; time \frac{(N-1)T}{N} \middle| I_B \; at\; 0\right)$$

So we may finally write,

$$Prob^C('survival', N) = \approx (1-\epsilon)^{N+1} \cos^2\left(B^C(0,T)\right) + \epsilon(1-\epsilon)^N \sin^2\left(B^C\left(\frac{(N-1)T}{N}, T\right)\right) \cos^2\left(B^C\left(0, \frac{(N-1)T}{N}\right)\right)$$

Additional details on the experimental methods.

Block	Evidence	Relative Strength, a_i	S.D.
1	Dixon was successful in his career and had recently been promoted.	0.92	0.49
	Dixon had arranged a number of social engagements for the week after his death.	0.83	0.48
	Dixon had no history of depression or related conditions.	0.94	0.48
2	Dixon was engaged to be married.	0.89	0.49
	One of Smith's previous housemates reported that Smith made him feel threatened.	1.15	0.50
	Friends and colleagues reported that Dixon did not seem obviously stressed or depressed in the days leading up to his death.	0.90	0.48
3	Neighbours reported overhearing Dixon and Smith engaged in heated conversations on the evening before Dixon's death.	1.25	0.43
	Dixon appeared to have a large quantity of savings.	0.70	0.46
	Smith had a previous conviction for assault.	1.22	0.44
4	Smith's fingerprints were found on the bottle of liquor, although it was impossible to tell whether these were recent.	1.01	0.53
	The addition of the sleeping pills to the liquor was unlikely to have altered its taste.	0.92	0.51
	The local pharmacist testified that Smith had bought the sleeping pills in his pharmacy recently after complaining of insomnia.	1.29	0.48

Table S1. The 12 pieces of evidence suggesting that Smith is guilty, with average relative strengths and standard deviations. This data was based on participants' judgments about the strength of evidence, as collected at the end of Experiments 1, 2. The average relative strength of evidence in blocks 1,2,3 and 4 is 0.90, 0.98, 1.06 and 1.07 respectively.

Details of the Bayesian Analyses

The computations of BIC and Bayes Factors were carried out following Jarosz and Wiley (22). In particular, the BIC was estimated from the R^2 via,

$$BIC = n * \ln(1 - R^2) + k * \ln(n)$$

Where k is the number of free parameters and n is the sample size. The Bayes factors were then computed in the usual way,

$$BF_{QB} = e^{\Delta BIC_{QB}/2}$$

where $\Delta BIC_{QB} = BIC_Q - BIC_B$ is the difference in BIC values for the Quantum and Bayesian models.

Additional references for Supplementary Materials

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