



Research

Cite this article: Yearsley JM, Pothos EM.

2016 Zeno's paradox in decision-making.

Proc. R. Soc. B **283**: 20160291.

<http://dx.doi.org/10.1098/rspb.2016.0291>

Received: 8 February 2016

Accepted: 14 March 2016

Subject Areas:

cognition

Keywords:

decision-making, opinion change,
constructive influences, quantum theory

Author for correspondence:

James M. Yearsley

e-mail: james.m.yearsley@vanderbilt.edu

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2016.0291> or via <http://rspb.royalsocietypublishing.org>.

Zeno's paradox in decision-making

James M. Yearsley¹ and Emmanuel M. Pothos²

¹Department of Psychology, Vanderbilt University, Nashville, TN, USA

²Department of Psychology, City University London, London, EC1V 0HB, UK

Classical probability theory has been influential in modelling 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 (QT). One unique feature of QT is the collapse postulate, which entails that measurements (or in decision-making, judgements) 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 judgements. 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.

1. 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 behaviour. 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 modelling.

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 judgements 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 judgements. From the QT model, we extract the surprising prediction of a QZ effect when intermediate judgements are made and contrast this with the prediction of the Bayesian model.

2. 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 two-dimensional 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 $N \geq 1$ judgements at times $T/N, 2T/N \dots T$ is (assuming a typical time independent Hamiltonian, all derivations in the electronic supplementary material):

$$\text{Prob}\left(I \text{ at time } \frac{T}{N} \wedge I \text{ at } \frac{2T}{N} \wedge \dots\right) = \cos^{2N}\left(\frac{\gamma}{N}\right). \quad (2.1)$$

Here γ is a dimensionless constant that encodes the effect of the evidence in the absence of intermediate judgements. As the number of measurements, N , increases, there is a decreasing probability that the system will change from I to G . As $N \rightarrow \infty$, 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].)

3. The quantum model

The derivation leading to equation (2.1) 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 $N \rightarrow \infty$, 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 modelled by a time independent unitary evolution. For the situation of interest, we may still assume the dynamics are approximately unitary (see the electronic 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 electronic 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 represents 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 approximately a_1 the second is given weight approximately $a_2 e^{-\beta}$, and so on.

Because 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 judgements 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 electronic supplementary material.

Using the above, we can show that:

$$\text{Prob}(I \text{ at } t | I \text{ at } 0) = (1 - \epsilon)^2 \cos^2(B(0, t)) + \epsilon(1 - \epsilon) \sin^2(B(0, t)). \quad (3.1)$$

The above equation 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 judgements).

We can also use equation (3.1), together with some assumptions about the way judgements 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 judgements the QT and Bayesian models will coincide. This means that we can use data obtained in the absence of any intermediate judgements to fix all the parameters in both the QT and Bayesian models. Our central predictions, of the specific way in which intermediate judgements affect opinion change, will therefore be parameter free.

4. 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 judgements increases. In the electronic supplementary material, we compute the probability of judging innocent at each of the intermediate judgements and the final one (N in total), given an initial innocence judgement. By analogy with the physics case, this can be called survival probability [14]. The result is

$$\begin{aligned} \text{Prob}^Q(\text{'survival'}, N) &= \text{Prob}\left(I \text{ at } \frac{T}{N} \wedge I \text{ at } \frac{2T}{N} \wedge \dots \wedge I \text{ at } T\right) \\ &= (1 - \epsilon)^{N+1} \prod_{i=0}^{N-1} \cos^2\left(B\left(\frac{iT}{N}, \frac{(i+1)T}{N}\right)\right) \\ &\quad + \epsilon(1 - \epsilon)^N \sin^2\left(B\left(\frac{(N-1)T}{N}, T\right)\right) \\ &\quad \times \prod_{i=0}^{N-2} \cos^2\left(B\left(\frac{iT}{N}, \frac{(i+1)T}{N}\right)\right) + O(\epsilon^2). \end{aligned} \quad (4.1)$$

The first term in this expression corresponds to the probability that the cognitive state is always consistent with innocent, and all the judgements reflect this. The second term corresponds to the possibility that the state changes between the second to last and final judgements, but the participant nevertheless responds 'innocent' due to the imperfect measurements. Further terms would correspond to more judgements not matching the cognitive state, or to the state changing back from innocent to guilty, these terms are negligible compared with those included in equation (4.1). If $\epsilon = 0$, $\beta = 0$ and the a_i 's are equal then equation (4.1) reduces to equation (2.1).

5. Constructing a matched Bayesian model

The QT model assumes that evidence changes the opinion state (as determined by equation (3.1)), that judgements may be imperfect and that judgements 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 equation (4.1); the survival probability after T pieces of evidence have been presented, given that N judgements have been made. This is

$$\text{Prob}^C(\text{'survival'}, N) = \text{Prob}\left(I_R \text{ at time } T \wedge I_R \text{ at time } \frac{(N-1)T}{N} \wedge \dots \wedge I_R \text{ at time } \frac{T}{N} \mid I_R \text{ at } 0\right).$$

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

As already noted, because equation (3.1) does not involve any intermediate judgements it may be interpreted

classically. We can therefore read off

$$\begin{aligned} \text{Prob}(I_B \text{ at time } t \mid I_B \text{ at time } 0) &= \cos^2(B(t, 0)), \\ \text{Prob}(G_B \text{ at time } t \mid I_B \text{ at time } 0) &= \sin^2(B(t, 0)), \\ \text{Prob}(I_R \text{ at time } t \mid I_B \text{ at time } t) &= (1 - \epsilon), \\ \text{Prob}(G_R \text{ at time } t \mid I_B \text{ at time } t) &= \epsilon \\ \text{Prob}(G_R \text{ at time } t \mid G_B \text{ at time } t) &= (1 - \epsilon), \\ \text{Prob}(I_R \text{ at time } t \mid G_B \text{ at time } t) &= \epsilon, \end{aligned}$$

(since the probabilities for judgements given cognitive states do not depend on the time, we may denote them simply as $\text{Prob}(I_R \mid I_B)$, 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 judgements.

When there are intermediate judgements 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 judgement rather than on the total time, effectively being 'reset' after every judgement. This is very natural from a QT perspective, however, the judgements 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 judgements 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 judgements. In particular, this means fitting either function to the data in the absence of intermediate judgements 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 electronic supplementary material)

$$\begin{aligned} \text{Prob}^C(\text{'survival'}, N) &= \text{Prob}\left(I_R \text{ at time } T \wedge I_R \text{ at time } \frac{(N-1)T}{N} \wedge \dots \wedge I_R \text{ at time } \frac{T}{N} \mid I_R \text{ at } 0\right) \\ &= (1 - \epsilon)^{N+1} \cos^2(B^C(0, T)) \\ &\quad + \epsilon(1 - \epsilon)^N \sin^2\left(B^C\left(\frac{(N-1)T}{N}, T\right)\right) \\ &\quad \times \cos^2\left(B^C\left(0, \frac{(N-1)T}{N}\right)\right) + O(\epsilon^2). \end{aligned} \quad (5.1)$$

The main feature of the Bayesian prediction is a reduction of survival probability with more intermediate judgements, because of a probability of error at each judgement. This contrasts sharply with the QT prediction, equation (4.1). 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 judgements. Would it be possible to include such influences? One way to do this might be to regard the memory of having made a previous judgement of guilt/innocence as additional evidence in favour 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.

6. Experimental investigation

(a) Participants

We ran the same experiment twice (Experiments 1 and 2), with different samples, solely as a replication exercise. Thus, we describe the two experiments together. For Experiment 1, we recruited 450 experimentally naive 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 naive participants from CrowdFlower. Participants were 39% male and 61% female (less than 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 min; Amazon Turk participants were paid \$0.50 and CrowdFlower participants \$1.00.

(b) 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 (adapted from [21] p. 287):

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 judgement 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 (electronic supplementary material, 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 judgement, but were also asked to make intermediate judgements (and justify their responses). These intermediate judgements 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 judgements 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 (electronic supplementary material, table S1).

(c) Results and model fits

Empirical assessment involved two steps. First, without intermediate judgements (i.e. at the first judgement made after having seen some evidence) the data are classical and simply informs us how opinion changes with evidence. Using equation (3.1), we can determine ϵ and $B(t_m, t_n)$ i.e. the

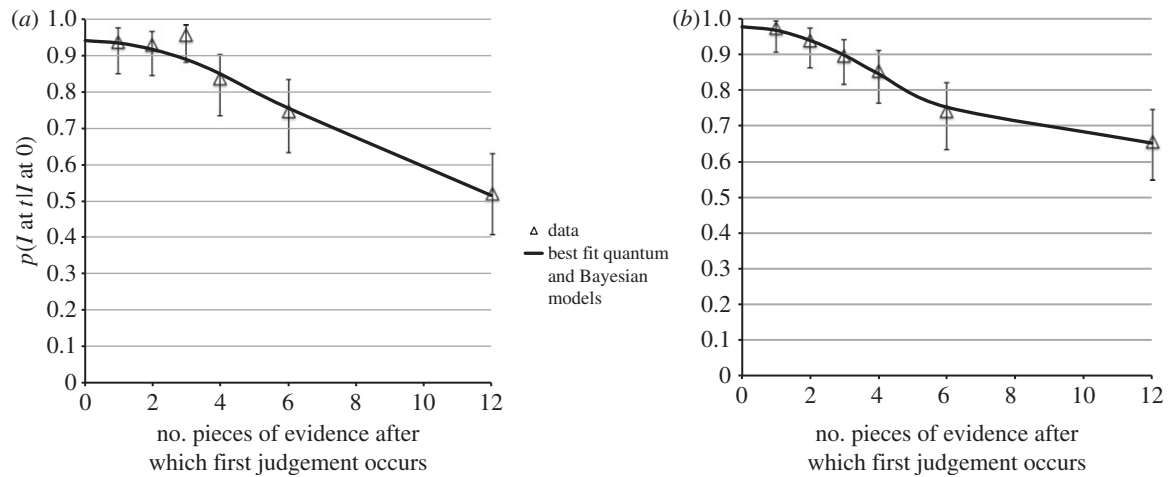


Figure 1. Setting the parameters (opinion change without intermediate judgements): $\text{Prob}(I \text{ at } t|I \text{ at } 0)$, for the first judgement a participant made, after having seen different numbers of pieces of evidence. (a) Experiment 1 (Amazon Turk). Note the obvious outlier at three pieces 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.

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 judgements produce the QZ effect (a slowing down of opinion change, as predicted by the QT model, equation (4.1)) or not (in which case the Bayesian model should fit better). The predictions about intermediate judgements 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 the electronic supplementary material on fixing the parameters; electronic supplementary material, 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 equation (3.1) and the data. For Experiment 1, and considering the $t = 3$ data point an outlier, the best fit for equation (3.1) is obtained with $\alpha = 0.091$, $\beta = 0.10$ and $\epsilon = 0.030$, giving an R^2 of 0.996 and a BIC of -27.8 . For Experiment 2, the best-fit parameters are $\alpha = 0.114$, $\beta = 0.285$ 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 , $\text{Prob}(I \text{ at } t|I \text{ at } 0)$ is nonlinear 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 , $\text{Prob}(I \text{ at } t|I \text{ at } 0)$ 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 $\text{Prob}(I \text{ at } t|I \text{ at } 0)$ eventually becomes linear in figure 1b. Also, more participants gave an initial judgement of 'guilty' in Experiment 2, compared to Experiment 1 (5% versus 11%). Despite distinct behavioural patterns across Experiments 1, 2, equation (3.1) 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 equations (4.1) and (5.1) to compute survival probabilities, for different numbers of intermediate judgements.

Empirical results for $\text{Prob}(\text{'survival'}, N)$ clearly favour 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 judgements, because of a probability of error at each judgement. For the QT model, in Experiment 1, we have a clear QZ effect, as survival probability generally increases with N . In Experiment 2, behaviour shows a tension between diminishing returns and QZ. With one intermediate judgement, the resetting of diminishing returns means that later pieces of evidence are weighted more strongly than in the case of no intermediate judgements, hence the dip in survival probability. With more intermediate judgements, 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 judgements.

There is an alternative test of the QT versus Bayesian models. We can employ equations (4.1) and (5.1) to compute survival probabilities for the condition where there is a judgement after every piece of evidence (number of pieces of evidence presented T , and number of judgements N , vary, but T/N fixed to 1). Again, the data clearly favour 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.

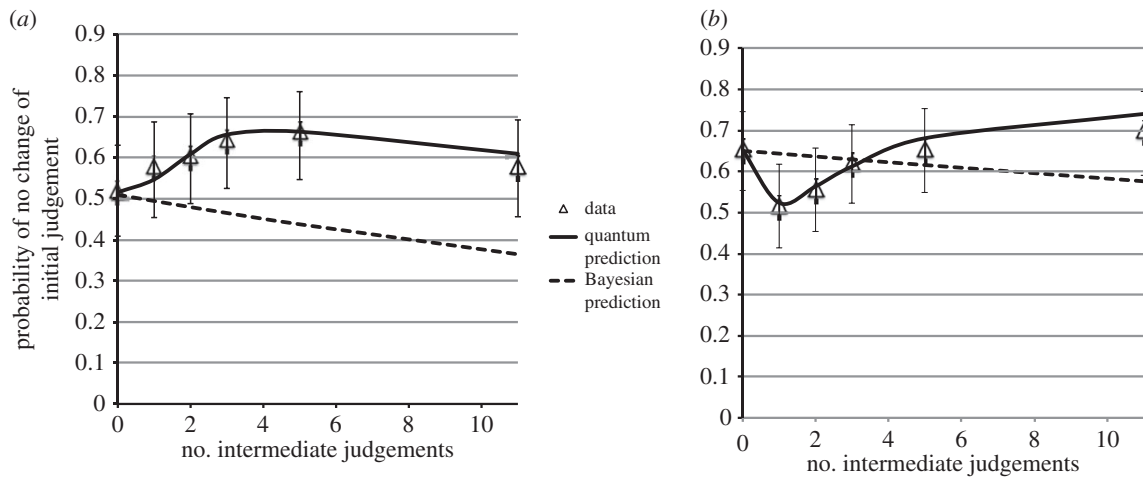


Figure 2. Evaluating the models: survival probability for N intermediate judgements, for the QT and Bayesian models, against empirical results ((a) Experiment 1, $N = 75, 71, 73, 70, 71, 64$, for each data point; (b) Experiment 2, $N = 90, 89, 88, 95, 81, 73$). Data points are participant averages and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

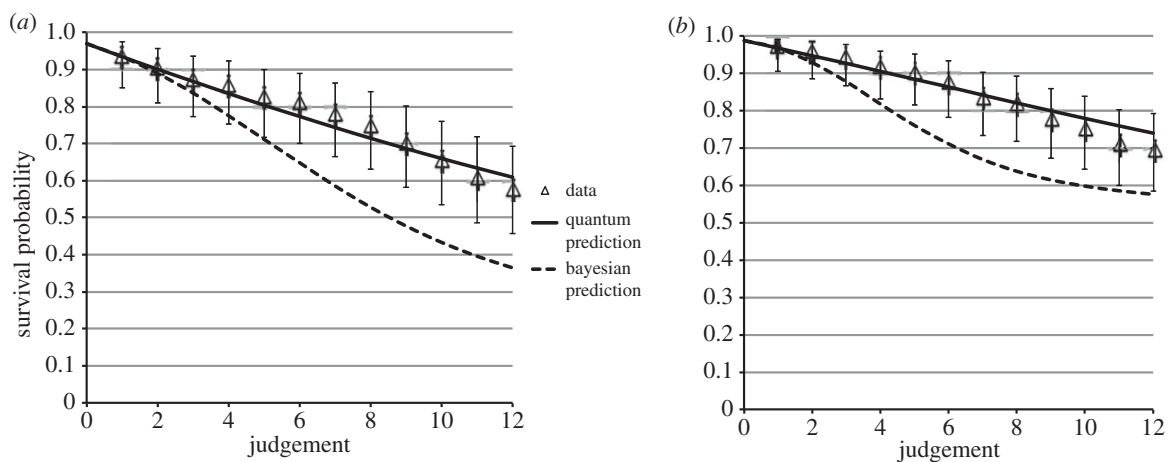


Figure 3. Evaluating the models: survival probability after each judgement, for the condition with 12 judgements ((a) Experiment 1, $N = 64$ for all data points; (b) Experiment 2, $N = 73$ for all data points). Data points are participant averages and error bars show 95% HDI of the posterior.

7. 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 judgements, 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 QT, the collapse postulate and *not* from parameter fixing. Our results show that decision theory needs to incorporate opinion influences from judgements. 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, behaviours paradoxical from Bayesian perspectives have often been interpreted as boundaries in the applicability of probabilistic modelling. 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 programme 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 also to non-humans. 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 behaviour seen in humans. This suggests perhaps judgements 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.

Data accessibility. Full datasets for the experiments reported in this paper are available via Dryad, <http://dx.doi.org/10.5061/dryad.n0k69>.

Authors' contributions. J.M.Y. conceived the project, developed the mathematical formalism, designed the experimental tasks, collected the data and analysed the results. E.M.P. helped develop the project and model, and contributed to the experimental design and the analysis of the results. Both authors worked equally on the manuscript.

Competing interests. We declare we have no competing interests.

Funding. E.M.P. and J.M.Y. were supported by Leverhulme Trust grant RPG-2013-00. Further, J.M.Y. was supported by an NSF grant SES-1556415 and E.M.P. was supported by Air Force Office of Scientific Research (AFOSR), Air Force Material Command, USAF, grants FA 8655-13-1-3044.

Acknowledgements. The US Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon. We would like to thank Jerome Busemeyer, David Lussseau and Jenifer Trueblood for useful comments and Thomas Zentall for bringing reference [24] to our attention.

References

- Anderson JR. 1990 *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Tenenbaum JB, Kemp C, Griffiths TL, Goodman N. 2011 How to grow a mind: statistics, structure, and abstraction. *Science* **331**, 1279–1285. (doi:10.1126/science.1192788)
- Tversky A, Kahneman D. 1983 Extensional versus intuitive reasoning: the conjunctive fallacy in probability judgment. *Psychol. Rev.* **90**, 293–315. (doi:10.1037/0033-295X.90.4.293)
- Hayden BY, Heilbronner SR, Nair AC, Platt ML. 2008 Cognitive influences on risk-seeking by rhesus macaques. *Judgment Decis. Mak.* **3**, 389–395.
- Aerts D, Aerts S. 1995 Applications of quantum statistics in psychological studies of decision processes. *Found. Sci.* **1**, 85–97. (doi:10.1007/BF00208726)
- Busemeyer JR, Bruza P. 2011 *Quantum models of cognition and decision making*. Cambridge, UK: Cambridge University Press.
- Haven E, Khrennikov A. 2013 *Quantum social science*. Cambridge, UK: Cambridge University Press.
- Pothos EM, Busemeyer JR. 2013 Can quantum probability provide a new direction for cognitive modeling? *Behav. Brain Sci.* **36**, 255–327. (doi:10.1017/S0140525X12001525)
- Wang Z, Solloway T, Shiffrin RM, Busemeyer JR. 2014 Context effects produced by question orders reveal quantum nature of human judgments. *Proc. Natl Acad. Sci. USA* **111**, 9431–9436. (doi:10.1073/pnas.1407756111)
- Pothos EM, Busemeyer JR. 2009 A quantum probability explanation for violations of 'rational' decision theory. *Proc. R. Soc. B* **276**, 2171–2178. (doi:10.1098/rspb.2009.0121)
- Hammeroff SR. 1998 Quantum computation in brain microtubules? The Penrose–Hammeroff 'Orch OR' model of consciousness. *Phil. Trans. R. Soc. A* **356**, 1869–1896. (doi:10.1098/rsta.1998.0254)
- Schwarz N. 2007 Attitude construction: evaluation in context. *Soc. Cogn.* **25**, 638–656. (doi:10.1521/soco.2007.25.5.638)
- White LC, Pothos EM, Busemeyer JR. 2014 Sometimes it does hurt to ask: the constructive role of articulating impressions. *Cognition* **133**, 48–64. (doi:10.1016/j.cognition.2014.05.015)
- Misra B, Sudarshan ECG. 1977 The Zeno's paradox in quantum theory. *J. Math. Phys.* **18**, 756–763. (doi:10.1063/1.523304)
- Atmanspacher H, Filk T, Romer H. 2004 Quantum Zeno features of bistable perception. *Biol. Cybern.* **90**, 33–40. (doi:10.1007/s00422-003-0436-4)
- Huggett N. 2010 Zeno's paradoxes. In *The Stanford encyclopedia of philosophy (Winter 2010 edition)*, (ed. EN Zalta). See plato.stanford.edu/archives/win2010/entries/paradox-zeno/.
- Anastopoulos C, Savvidou N. 2006 Time-of-arrival probabilities and quantum measurements. *J. Math. Phys.* **47**, 122106. (doi:10.1063/1.2399085)
- Yearsley JM, Pothos EM. 2014 Challenging the classical notion of time in cognition: a quantum perspective. *Proc. R. Soc. B* **281**, 20133056. (doi:10.1098/rspb.2013.3056)
- Isham C. 1995 *Lectures on quantum theory*. London, UK: Imperial College Press.
- Trueblood JS, Busemeyer JR. 2013 A comparison of the belief-adjustment model and the quantum inference model as explanations of order effects in human inference. *Cogn. Sci.* **35**, 1518–1552. (doi:10.1111/j.1551-6709.2011.01197.x)
- Tetlock PE. 1983 Accountability and the perseverance of first impressions. *Soc. Psychol. Q.* **46**, 285–292. (doi:10.2307/3033716)
- Jarosz AF, Wiley J. 2014 What are the odds? A practical guide to computing and reporting Bayes factors. *J. Probl. Solving* **7**, 2–9. (doi:10.7771/1932-6246.1167)
- Kruschke JK, Vanpaemel W. 2015 Bayesian estimation in hierarchical models. In *The Oxford handbook of computational and mathematical psychology* (eds JR Busemeyer, Z Wang, JT Townsend, A Eidels), pp. 279–299. Oxford, UK: Oxford University Press.
- Oaksford M, Chater N. 2009 Precipitous Bayesian rationality: the probabilistic approach to human reasoning. *Behav. Brain Sci.* **32**, 69–120. (doi:10.1017/S0140525X09000284)
- Herbranson WT, 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. (doi:10.1037/a0017703)