Pothos, E. M., Shiffrin, R. M. & Busemeyer, J. R. (2013). The dynamics of decision making when probabilities are vaguely specified. Journal of Mathematical Psychology, 59, pp. 6-17. doi: 10.1016/j.jmp.2013.09.001

The dynamics of decision making when probabilities are vaguely specified

Emmanuel M. Pothos[1]
Richard M. Shiffrin[2]
Jerome R. Busemeyer[2]

Submitted for the special issue honoring the contributions of William K. Estes.

Correspondence/ affiliations:
1. City University London, Department of Psychology, London, EC1V 0HB, UK; e.m.pothos@gmail.com
2. Indiana University, Department of Psychological and Brain Sciences, 10th St., Bloomington, IN 47405, USA; shiffrin@indiana.edu, jbusemey@indiana.edu

Word count: 13,418; Running head: Quantum vagueness

Abstract

Consider a multi-trial game with the goal of maximizing a quantity Q(N). At each trial N, the player doubles the accumulated quantity, unless the trial number is Y, in which case all is lost and the game ends. The expected quantity for the next trial will favor continuing play, as long as the probability that the next trial is Y is less than one half. Y is vaguely specified (e.g., someone is asked to fill a sheet of paper with digits, which are then permuted to produce Y). Conditional on reaching trial N, we argue that the probability that the next trial is Y is extremely small (much less than one half), and that this holds for any N. Thus, single trial reasoning recommends one should always play, but this guarantees eventual ruin in the game. It is necessary to stop, but how can a decision to stop on N be justified, and how can N be chosen? The paradox and the conflict between what seem to be two equally plausible lines of reasoning are caused by the vagueness in the specification of the critical trial Y. Many everyday reasoning situations involve analogous situations of vagueness, in specifying probabilities, values, and/or alternatives, whether in the context of sequential decisions or single decisions. We present a computational scheme for addressing the problem of vagueness in the above game, based on quantum probability theory. The key aspect of our proposal is the idea that the range of stopping rules can be represented as a superposition state, in which the player cannot be assumed to believe in any specific stopping rule. This scheme reveals certain interesting properties, regarding the dynamics of when to stop to play.

Keywords: vagueness, decision making, doubling game, quantum probability

Highlights:

*We consider decision making with vague probability information.

*We base our discussion on a finite version of the doubling game.

*We develop a quantum probability model for the dynamics of decision making in such a game.

*We present a way in which the model can be applied in scaled-down relevant empirical situations.

*We report empirical results, which allow a preliminary assessment of the methods and manipulations.

1. Introduction

The work of William Estes has had a lasting and profound influence in the development of psychological theory. One important contribution was the instigation of a revolutionary transition from descriptive to quantitative theories, the latter presented precisely in mathematical language. A mathematical formalization of a psychological process can provide a framework within which to study what is possible and what is not, identify key theoretical issues, and provide a guide for future empirical exploration. Our aim fits such goals: we study an important theoretical problem, in the psychology of decision making, and provide a computational framework for the assumed underlying cognitive processes. We then explore the properties of the computational framework, derive an empirical prediction, and experimentally test this.

Consider a 'doubling game', in which a player starts with 1 unit. On each trial, the player can choose to either stop playing and take home her winnings or double her accumulated units. However, if she doubles on trial Y, she loses all and the game ends. The number of Y has been vaguely specified, e.g., a person filled a sheet with random digits, permuted them, and so produced Y (since it is impossible to write an infinite number of digits on a sheet of paper, the number Y is finite). To justify stopping on trial N, the player must think that Prob(N=Y) is at least ½. But this is always unreasonable. The player can attempt to guess the probability distribution for Y (i.e. a Bayesian prior). However, having successfully reached trial N-1 without losing, she will surely form a posterior distribution that predicts considerable probability for numbers above N. Why? Because, seeing N-1 ensures that the number of digits written down were sufficiently numerous to generate N-1. Knowing this makes it likely that numbers higher than N-1 will be produced. Thus, the paradox in the doubling game is caused by the vague specification of Y (e.g., Bonini et al., 1999). Note that a prior could be chosen that will delay the stopping decision to such a large trial number (e.g. 100**(100**100)) that ruin would certainly occur, but this prior is not helpful. Whichever way one attempts to specify a prior on Y, we suggest that this will not resolve the paradox.

The prescription to continue playing is paradoxical, since eventually Y will be reached. This game is a finite version of the one in St. Petersburg's paradox, in which a fair coin is tossed on every trial, tripling the accumulated payoff on obtaining e.g. heads. The problem is general, e.g., it can be recast in a single trial version and can involve payoff functions, which recommend playing, regardless of the value of Prob(next guess=Y) and regardless of utility functions incorporating e.g. loss aversion or diminishing returns (the latter was Bernoulli's approach in dealing with St. Petersburg's paradox).

We think that there does not now exist, and may never exist, a normative theory of rationality that could be used to provide an ideal basis for decision making. Rather, for a particular problem, people explore different perspectives of reasoning, in an attempt to identify the best course of action. When the different perspectives converge, it becomes easy to resolve the decision making problem. When they diverge, then inevitably the decision making process is more difficult. In the doubling game, a local trial perspective mandates doubling. But is this reasoning correct? A player will attempt to verify correctness by trying out other perspectives. One is a global perspective and leads the player to ask whether he will gain if he continues to play indefinitely. However, since this guarantees eventual ruin, the global perspective recommends stopping at some trial. Thus, here, the two different perspectives, local and global, are in conflict.

How does one deal with conflicting, and yet plausible, lines of reasoning (or goals, or perspectives) for a problem? Notwithstanding the normative problems, we expect that human decision makers in a doubling game will eventually stop playing. We develop a descriptive model of decision making in a doubling game, so as to as to provide a basis for further systematic study of this behavior. We propose to approach the problem using quantum probability (QP) theory, by which we mean the probabilistic calculus of quantum mechanics, not specifically tied to physics (e.g., Aerts & Gabora, 2005; Atmanspacher, Romer, & Wallach, 2002). QP theory is basically a way to quantify uncertainty, which is alternative to classical, Bayesian probability (CP) theory. Predictions from QP and CP models sometimes converge, but QP theory has some properties that distinguish it from CP theory. For example, in QP theory probabilistic assessment can be strongly order and context dependence and superposition states exist, for which it is not possible to make precise statements regarding certain questions. Thus, QP models have usually been proposed to cover empirical situations for which it has been difficult to develop satisfactory CP models (e.g., Aerts & Gabora, 2005; Bruza et al., 2009; Busemeyer et al., 2011; Busemeyer & Bruza, 2012; Khrennikov, 2010; Pothos & Busemeyer, 2009; Trueblood & Busemeyer, 2011; Wang & Busemeyer, in press).

One motivation for adopting QP theory here is that, traditionally, QP cognitive models have fared well in situations which appear paradoxical from a CP perspective (Busemeyer & Bruza, 2012; Pothos & Busemeyer, in press; Wang et al., in press). For example, in a QP model many different lines of reasoning can co-exist in a superposition state, until a judgment is made, at which point one approach emerges to govern the decision. In a doubling game, belief regarding the next stopping rule could simultaneously incorporate biases both about large stopping rules (favoring a local perspective of the game, according to which it is advantageous to continue playing for as long as possible) and smaller stopping rules (favoring the global perspective of the game and a bias to stop). However, as we discuss (General discussion), we think that many of the ideas in the present QP model can be implemented classically. Another motivation is that, because most decision making models have, until recently, focused on traditional methods (based on traditional probability measures), it is interesting to provide some groundwork for how a QP theory approach can handle decision making behavior, since, overall, both CP theory and the QP theory approaches have merits and demerits and both are worth implementing and testing in empirical research. A final motivation is that the QP model enabled the extraction of analytical solutions fairly easily.

The specification of the model allows us to consider aspects of a doubling game, which may impact on behavior. Accordingly, we next develop a simple empirical task, which tests a corresponding prediction. We note that such tasks are likely to involve considerable conceptual and methodological challenges, but their development is essential in order to appreciate how naïve observers cope with problems that are problematic from a normative perspective.

Intuition regarding behavior in the doubling game enables the following assumptions, to guide model construction. First, there is a ***bias*** to continue playing on any particular trial. This reflects the local perspective of the game, since the game is set up in such a way that it always makes sense to continue playing. For small trial numbers, this is because of the extremely small probability that such numbers could be Y. For larger trial numbers, as argued, if a player has already observed N-1, she is unlikely to believe that Y would be N. Second, a player creates a guess for when to stop. Regardless of how such a guess is created, we assume that, when the player reaches this stopping rule, she stops playing. Finally, as the game proceeds, and obviously as long as the game has not yet ended in ruin, we assume that the player might decide to change her stopping rule. For example, if the player had decided to stop after 10 trials, and she has already won in the first 9 trials, would this not make it more likely that she will continue playing? Then, a cognitive model involves a specification of how the evidence from winning successive trials impacts on confidence for the (latest) stopping rule and potentially leads to its revision.

2. Model processes and results

We first describe the model mostly conceptually and then develop the mathematical details. In the mathematical details below, a lot of the arithmetic is simple linear matrix algebra; where relevant, we make some corresponding notes, which will hopefully make the material more accessible.

We represent information about the *next* guess regarding the appropriate stopping rule, with a vector $|\psi\rangle$, called the state vector (because it describes the state), in a complex Hilbert space, such that each one of its axis is a possible stopping rule. Hilbert spaces are the vector spaces employed QP theory; they are complex vector spaces, with some convergence properties as well. Complex numbers (i.e., numbers of the form $a + i \cdot b$, $i = \sqrt{-1}$) are employed in QP theory, as a mathematical convenience. All the aspects of a system that can be observed, called, surprisingly enough, observables, are guaranteed to have real values. Also, the set of axes we employ to characterize a Hilbert space are more accurately described as basis vectors, i.e., a set of orthonormal vectors, such that any vector in the space can be expressed as a linear combination of them. Finally, we employ the Dirac bra – ket notation, in which $|\psi\rangle$ indicates a column vector and $\langle\psi|$ denotes the complex conjugate of the transpose of $|\psi\rangle$ (i.e., a row vector). For example, let us say that the possible basis vectors are $|1\rangle, |2\rangle, \ldots |Max\rangle$, each one indicating belief that a particular stopping rule is appropriate. Then, if $|\psi\rangle = |10\rangle$, this means that the player is certain he ought to stop at trial 10. In general, $|\psi\rangle$ would incorporate belief about a range of possible stopping rules, so that, for example, we would have $|\psi\rangle = |1\rangle + |2\rangle + \cdots |Max\rangle$. Assuming equal belief in all stopping rules, each basis vector needs be multiplied by a normalization factor of $\frac{1}{\sqrt{max}}$, which we do not show for notational convenience. The factor corresponding to each basis vector is called the amplitude and it reflects probability weight. For example, say the amplitude associated with the $|2\rangle$ basis vector is a, where a could be a complex number. Then, the probability that a measurement will show $|\psi\rangle = |2\rangle$ is $a \cdot a^* = |a|^2$. Note also that we have to limit the overall dimensionality of the space, i.e., there is a maximum stopping rule, *Max*.

When $|\psi\rangle = |1\rangle + |2\rangle + \cdots |Max\rangle$, we have what is called a superposition state. Even though such a state looks like a classical linear combination, in QP theory, in such a state, the player cannot be assumed to believe in any stopping rule. An analogy with physics is useful. Suppose that a state vector describes the position of a particle, which, in this simplified example, can be either left or right. A classical state vector could be $p_1 Left + p_2 Right$, whereby p_1 is the probability that the particle is on the left and p_2 on the right. Crucially, the particle is assumed to be definitely either left or right, it is just that we are not sure of which possibility is true. By contrast, a quantum state vector $\sqrt{p_1} \cdot |Left\rangle + \sqrt{p_2} \cdot |Right\rangle$ expresses the uncertainty in the *potentiality* of whether the particle will be left or right. However, while the state is in a superposition the particle cannot be assumed to be either left or right (it cannot be assumed to be at both locations concurrently as well)! That this has to be the case in QP theory is a requirement from a famous mathematical result (the Kochen-Specker theorem; e.g., see Isham, 1989, or Hughes, 1989). In physical applications of QP theory, the implications of superposition have caused endless debate (a famous example is the so-called Schroedinger's cat). In psychology, perhaps the idea of superposition is easier to accept and indeed decision theorists sometimes do invoke the idea that there is no specific knowledge regarding belief, relevant to a decision/action (e.g., see Paulus & Yu, 2012) and that a 'measurement' may have constructive properties (e.g., Sharot et al., 2010).

Psychologically, a measurement could be a judgment, decision, or expression of preference. If the state is a superposition one, then a measurement will lead to a so-called 'collapse' of the superposition state into a definite state, for example, from $|\psi\rangle = |1\rangle + |2\rangle + \cdots |Max\rangle$ to say $|\psi\rangle = |2\rangle$.

As the player keeps winning, we assume that she revises her perception of what is the best stopping rule. After deciding on a particular stopping rule, the state vector would be set equal to the corresponding basis vector. But, the impact of winning on each trial is to undermine confidence in the latest guess for the stopping rule and increase the potentiality of higher stopping rules. This process is captured with a unitary matrix U. Unitary matrices in QP theory govern the dynamics of a quantum system. Presently, we need a U, such that the amplitude for all states which are lower than the latest stopping rule is zero and stays to zero, the amplitude for the state corresponding to the latest stopping rule is reduced with every trial, and the amplitude for all states higher than the latest stopping rule increases. If a state is expressed as $|\psi\rangle = a|1\rangle + b|2\rangle$, if the amplitude a is reduced and the amplitude b is increased, then the probabilities that a measurement will reveal $|\psi\rangle = |1\rangle$ versus $|\psi\rangle = |2\rangle$ change too (recall, the probability that the state becomes $|\psi\rangle = |1\rangle$ is $|a|^2 = a^* a$).

To illustrate, suppose that *Max*=10, that is, the absolute maximum number of trials is 10. Suppose also that the latest guess for a stopping rule is 4, that is, the player thinks that she should not play beyond the fourth trial. Figure 1 shows the dynamics of the model. Each sub-figure illustrates the probability (the modulus of each amplitude squared) for each state, that is, each possibility for the player's *next* stopping rule. Because the current stopping rule is 4, at time=0, all the probability is concentrated on this state, that is ($|\psi\rangle = |4\rangle$). At time=0, the player believes that her stopping rule should be 4, that is, that she should stop playing when she reaches the fourth trial. However, with every win confidence in this stopping rule is undermined, while preference for higher stopping rules increased. This means that at times>0, the state $|\psi\rangle$ expressing belief of what should be the next stopping rule becomes a superposition. Note that we assume that the player always plays on the basis of the latest stopping rule, and the distribution of amplitude in $|\psi\rangle$ just impacts on the possibilities for the next guess. In any case, as time increases the amplitude lost by $|4\rangle$ is shared by all higher states (this sharing of amplitude is equal for all for all higher states; technically, this is the simplest way to implement the model). Thus, while any specific higher state will still be relatively improbable, collectively eventually they become more probable than the state corresponding to the latest guess, $|4\rangle$. Note also that the amplitude of all states lower than $|4\rangle$ starts at 0 and stays at 0, since what would be the logic in deciding to adopt a lower stopping rule? At about t=0.6 the probability for $|4\rangle$ drops to 0 and at that time point the probability for all higher states is maximum (time units are arbitrary; if correspondence with real time is required, then a units constant needs be introduced). After this minimum/ maximum, the changes in amplitudes start being reversed. Amplitude (and so probability) variation is periodic (this is, a common property of QP models; e.g., see also Busemeyer, Wang, & Townsend, 2006). The model is specified in such a way that any action relating to $|\psi\rangle$ is taken before the first maximum/ minimum is reached.

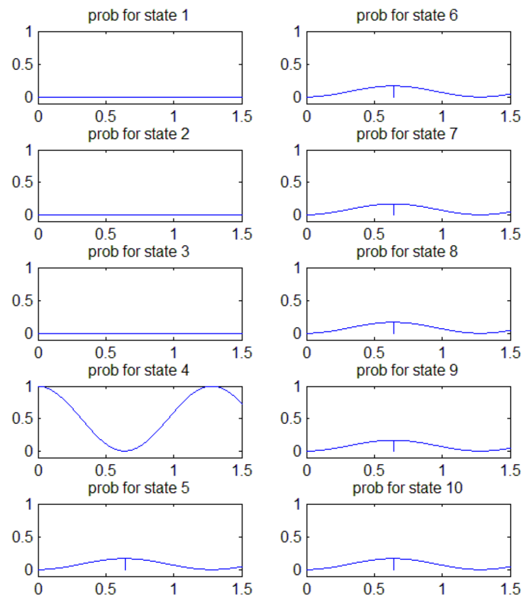

Figure 1. An illustration of the dynamics of the QP theory model, for a situation when the latest stopping rule is 4 and the maximum possible stopping rule is 10. Vertical axes show probability. Horizontal axes show time.

We modify the above reasoning for the first stopping rule. This would be the guess, prior to having any experience with the task. We suggest that when such a guess is formulated, the state $|\psi\rangle$ just after the guess is not restricted to the basis state of the guess value, but rather is a maximally vague superposition state. This assumption is reasonable, in that a player is unlikely to have any confidence in the first guess for the stopping rule, over and above alternative possibilities. An example of the model dynamics in this case is shown in Figure 2. In this case, all states start with the same amplitude. Since there are 10 states, then the amplitude of each state is $1/\sqrt{10}$ and so the initial probability of each state is $1/10$. Let us assume that the initial (wild) guess for a stopping rule is 4. Then, with increasing time, the amplitude for state $|4\rangle$ and all lower states is reduced and the amplitude for all higher states is increased.

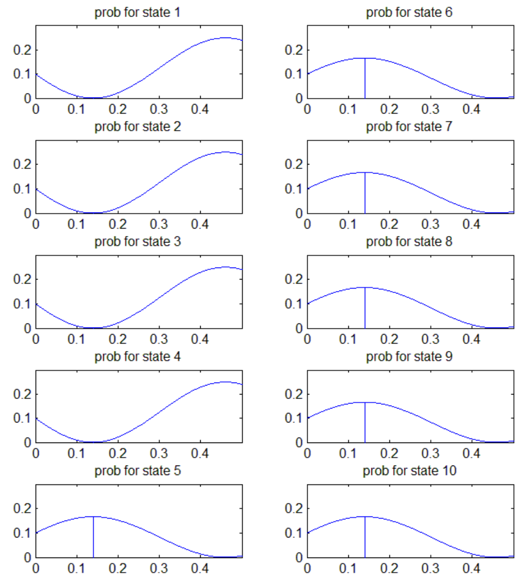

Figure 2. An illustration of the dynamics of the QP theory model, for the first guess for a stopping rule, when this guess is 4 and the maximum possible stopping rule is 10. Vertical axes show probability. Horizontal axes show time.

To summarize, $|\psi\rangle$ contains all the information regarding beliefs about a next guess for the stopping rule. When a stopping rule is decided, the state $|\psi\rangle$ is set to the corresponding basis vector (unless, this is the very first guess). This means that, shortly after the decision for a particular stopping rule, if the player were asked to evaluate the evidence for possible stopping rules, he would re-select the stopping rule currently adopted. With time, probability weight in $|\psi\rangle$ is distributed across all possible stopping rules. In QP theory, when a system is described by a superposition state, to resolve this superposition and determine a specific value for the system, we need to perform a measurement, which reduces (collapses) the superposition state to a particular basis state. Which basis state the reduction occurs to is probabilistic and depends on the amplitude of each state. Note that the dynamics of QP theory cannot account for the transition from a superposition state to a basis state (Isham, 1989).

In the doubling game model, the state $|\psi\rangle$ will typically be a superposition state. Therefore, the player would be uncertain about the next guess (i.e., whether she should stick to the current stopping rule, or change to a higher one). We assume that at each trial, there is a probability that a reduction of $|\psi\rangle$ will occur , which means that $|\psi\rangle$ is set equal to a (probabilistically chosen) basis state. Because it is not possible to determine this probability within QP theory, we allow the probability of state reduction in each trial to depend on a parameter in the model, called $ExCo$ (expected collapses). This parameter is set so that the minimum number of expected state reductions before the first maximum/ minimum (in the probability associated with each state; for the states whose amplitude is increasing, the relevant turning point is the maximum and vice versa) is 1 and the maximum number is one per trial. Specifically, let tstep be the time advance associated with each trial (the size of tstep effectively determines the push away from the current stopping rule and in favor of higher stopping rules, with each trial), tstart the time point right after the adoption of a new stopping rule (this can be set to 0), and tmax the time point when the amplitudes for different possibilities for the next stopping rule reach their first maximum. Then, we can set the probability of a collapse of $|\psi\rangle$ on each trial to be equal to

$\frac{tstep}{tmax-tstart} \cdot ExCo$. Noting that there are $\frac{tmax-tstart}{tstep}$ trials between right after $|\psi\rangle$ has collapsed and the first maximum, then when $ExCo = 1$ we expect one collapse in that period. With a high enough value of $ExCo$, at most we can expect $\frac{tmax-tstart}{tstep}$ collapses (one per trial). Note, the value of tmax depends on both the highest possible stopping rule and the latest guess. Therefore, the $ExCo$ parameter can only be interpreted as a bias towards a greater or smaller number of collapses, rather than an absolute measure of how many collapses we expect.

So, what determines whether the model is likely to move away from the current guess to a new one? When a collapse of $|\psi\rangle$ takes place closer to tmax, this means that the probability weight for the latest guess for the stopping rule is close to 0 and the probability for higher guesses as high as possible. Therefore, the model is likely to pick a new, higher stopping rule. By contrast, when a collapse takes place close to tstart, this means that the probability for the latest guess is still high and the probability for higher guesses low. Therefore, the model is likely to reaffirm its current guess and not change. The more the model sticks to its latest guess for the stopping rule, the more likely it is that the latest guess will catch up with the current trial, at which point the player will stop playing. This is the way in which stopping behavior emerges from the present model. Clearly, the proposed formalism is not a normative solution to the doubling game and its paradox. However, it does provide a perspective for its resolution, at the level of psychological behavior.

We consider some patterns in the model's behavior. The higher the value of the $ExCo$ parameter (i.e., the higher the bias to collapse $|\psi\rangle$ at each time step or trial) and the smaller the length of each time step (i.e., the greater the number of trials between tstart and tmax), the more likely it is that $|\psi\rangle$ will collapse soon after the latest guess. This means that it is more likely to persevere with the latest guess for longer which, in turn, means that the player would be both more likely to stop playing sooner and be less flexible (in terms of changing his guess). By contrast, if there is a smaller $ExCo$ parameter and a larger time step, by the time $|\psi\rangle$ has collapsed, most of the weight would have been transferred to higher states, and this makes it more likely that the player will change his stopping rule to a higher one. Simply put, the *ExCo* parameter determines the number of times a player switches his stopping rule. As noted, the property of superposition implies that, once $|\psi\rangle$ does become a superposition, the state is not consistent with any specific possibility. From this fact, one could make the additional assumption that it is not possible to have insight about the accumulated evidence. Thus, the player will be content to continue playing on the basis of the latest stopping rule, even if the bulk of amplitude has been transferred to higher stopping rules. Only when $|\psi\rangle$ collapses, can we assume that the player employs up-to-date knowledge in favor of higher stopping rules.

We next describe simulations illustrating the function of the model (Table 1). Recall, the size of the time step determines the number of trials between tstart and tmax. Because on each trial there is a probability that $|\psi\rangle$ will collapse, the more the trials before tmax, the greater the likelihood that $|\psi\rangle$ will collapse sooner (at a time point closer to tstart) than later (at a time point closer to tmax). The easiest way to understand the behavior of the model is by fixing the time step parameter (0.001) and manipulating the collapse parameter ($ExCo$). The values in each row of Table 1 were determined by averaging results across 100 runs of the model. Note first that, as the collapse parameter increases, we end up with a higher number of collapses of the state vector $|\psi\rangle$. This is just an implication of what the collapse parameter does. In the second and third column we verify the key assumption regarding the function of the model. When there are fewer collapses, this means that $|\psi\rangle$ collapses closer to tmax, so that most of the amplitude would have been transferred to stopping rules higher than the latest one. This makes it more likely that the player will abandon the latest stopping rule and favor a higher one. Indeed, with fewer collapses, the model ends up preferring a higher stopping rule more frequently (i.e., we observe a greater number of switches in the guess). But, if the player keeps adopting higher stopping rules, then this makes it more difficult for the current trial to catch up with the latest stopping rule. In other words, it becomes less likely that, as the game progresses and the player goes through more trials, that he will reach the trial specified by his latest stopping rule – and so stop playing. By contrast, when the collapse parameter is high, $|\psi\rangle$ is likely to collapse closer to tstart, than tmax, which means that most of the probability weight would still be concentrated to the latest guess. Therefore, the player is less likely to change the guess for the stopping rule (low number of times guess switches) and he is also more likely to stop playing the game sooner (since, the more often he reaffirms the latest stopping rule, the more likely it is that the advancing trials will eventually reach the trial specified in the stopping rule, so that the game would stop). These expectations were confirmed in Table 1. Note that the model's dynamics are set up in such a way that stopping behavior can emerge, but is not required.

We illustrate, in Table 2, with a reduced set of simulations, the impact of reducing the time step, from 0.001 to 0.0001. A smaller time step means that, the transfer of amplitude from the current stopping rule to higher stopping rules, per trial, is also smaller. As can be seen, the trial on which the player is predicted to stop has been reduced, when using the smaller time step. Moreover, the number of switches from the latest guess for the stopping rules to higher stopping rules has also been reduced. With this lower time step, as the *ExCo* parameter increases, the player just stops on the basis of the first stopping rule specified. That is, the player perseveres with the initial guess for when he should stop playing the game. When the *ExCo* parameter is lower, there is a very slightly higher bias of switching stopping rule, but this is still low compared to the corresponding values with the higher time step.

The model is analytically presented in Appendix 1.

Table 1. Simulations illustrating the behavior of the model, with a time step of 0.001.

Collapse parameter	number of collapses	number of times guess switches	trial when game stopped	Number of times the game did not stop
2.5	13.65	4.35	917.43	7
5	25.27	4.02	910.58	16
10	56.18	3.49	865.64	4
15	84.52	2.54	716.41	4
20	124.03	2.20	766.96	2
25	155.30	1.82	678.62	2
30	175.35	1.74	685.12	0
35	212.71	1.62	668.51	0

Table 2. Simulations illustrating the behavior of the model, with a time step of 0.0001.

Collapse parameter	number of collapses	number of times guess	trial when game	Number of times the game did not stop

		switches	stopped	
2.5	2.40	1.13	599.99	1
10	6.76	1.17	594.90	0
20	13.12	1.12	539.61	0

3. Empirical illustration

The QP theory model is the first attempt to study the cognitive processes relevant to the doubling game and no doubt revisions will be necessary. Nonetheless, formalization has helped specify some of the factors which might influence behavior in the task. In the tradition of Estes's work, this is one way in which mathematical description can help psychological science, through its generative role regarding novel predictions about behavior.

Specifically, allowing more time per trial in the doubling game should lead to a higher overall stopping trial (that is, the trial when a participant decides to stop), and vice versa. Recall, the more the time per trial, the more amplitude transfer we have from the latest stopping rule to higher ones, per trial. Assuming that the probability of collapsing the state vector does not depend on the time per trial (rather, it depends on the number of trials; Appendix 1), and assuming that the thought process per trial does not extend beyond the first turning point for the amplitudes, with more time per trial, by the time the state vector collapses, more amplitude would have been transferred from the latest stopping rules to higher ones. This, in turn, makes it likely that, when a collapse occurs, a participant will adopt a higher stopping rule, instead of reaffirming the latest one. The adoption of higher stopping rules implies that a participant is likely to continue playing for longer. That this might be the case is a novel intuition regarding a doubling game task, enabled by the QP theory model. We next describe a simple laboratory version of the doubling game.

3.1 Participants and design

We recruited 39 participants, all students at the City University London, who took part for a small payment (£5). In addition, participants were told during the task that "good performance will be rewarded and you may end up earning a few more £, if you do well". The vagueness of these instructions was intentional, since more specific instructions may have altered the nature of the task. In practice, of the top performing participants (the ones who had not decided to stop playing by trial 75, see later), we randomly selected one and invited him/her to receive an additional £10. The main design of the experiment involved a between participants factor, relating to whether the time prescribed per trial was long or short.

3.2 Materials and procedure

The experiment was run on a computer. Participants were told that they would participate in a doubling game, which was described as a game played over several trials, in which they would start with a small amount of money and had to try to earn as much money as possible. The doubling game task was presented and participants were told of the 'secret' number Y, such that if the next trial N=Y, then they would lose all the money they had earned. Regarding the specification of Y, participants were told that "Someone was asked to choose a number, not too large, and this number is *Y*". It was noted that participants would not earn real money, but that good performance would be rewarded. They were finally told that they would go through a few practice trials, before the main task.

The trial structure for each practice trial was identical to that for the main task trials (to be described shortly), but for the fact that, after each practice trial, participants were given some additional commentary regarding the task. Specifically, as long as a participant decided to continue playing, after each practice trial, he/she was also told that, as they won the trial ($N<Y$), they doubled their money. Another purpose of the practice part of the experiment was to have a low *Y*, only 7. It was expected that participants, having reached trial 6, would continue playing, so that they would lose on trial 7. The reason for having a low *Y* was to give participants an example of losing a doubling game, so that, just curiosity of what happens when *Y* is reached, would be less of a factor in deciding to continue playing. Note that the practice trials possibly establish a prior for *Y*, even though there was nothing to indicate to participants that the *Y* in the practice game was related to the *Y* in the actual game. Equally, given that participants were brought to the lab for a 30 minutes session, it is likely that they would have an expectation that *Y* could not be 'too large' anyway.

After the practice trials, and assuming participants had no questions, the experiment continued with the set of experimental trials. Participants' starting amount of money was £0.0001. Given that, with each trial of continuing play, participants were meant to double their money, we restricted the total number of trials to 75. Participants playing all the trials would finish the game having earned £1,888,946,593 billion. Up to trial 44, the amount of money won was expressed in £, after trial 44 in billion £, to avoid having many digits on the screen. In the main task, there was no *Y*, that is, there was no trial at which participants would lose all the money they had gained. After trial 75, participants were simply told that the experiment had ended. So, one way in which the experiment would fail to work is if all participants continued playing beyond trial 75, but this was not the case. If participants decided to stop playing before trial 75, the computer programme prompted them to state (as a percentage) the probability that the next trial is *Y*. This manipulation was introduced as a check that participants were broadly behaving on the basis of expected value maximization. Given that winning doubles the accumulated amount, even for loss averse decision makers, it would only make sense for participants to stop playing, if they thought that the probability that the next trial is *Y* is high.

On each trial, at the top of the screen, participants saw information about the amount of money won so far and the current trial, for example, "Your current amount is £0.0002, this is trial 2". In the middle of the screen was the question "Do you want to continue playing? (Yes/ No)". In the bottom of the screen, participants saw a reminder about the structure of the game, that is, that if the next trial was less than *Y*, they would double their money, but if the next trial was equal to *Y* they would lose everything. Participants responded with a simple yes or no answer, by pressing an appropriately labeled key.

For the first 11 participants, the instructions stated that participants would have as much time as they wanted to respond, but that they should try to not take too long. Accordingly, participants had an unlimited amount of time, per trial, to provide a response. The results from these participants were employed so as to specify appropriate short or long response times, which would in turn inform the main between-participants manipulation. We briefly describe these results here, as they inform the experimental design. For each participant we computed the average response time, across all the played trials. These averages ranged from 768ms to 5705ms, thus, it is clear that there is considerable variability in the speed with which the relevant decisions are made. From these averages, our aim was to identify appropriate time limits for the Short, Long conditions in the experiment. A complication is that, if the allowed time for the Short condition trials were too short, then participants would be likely to not respond within the prescribed time. If the allowed time for the Long condition trials were too long, participants might get bored (and, e.g., be tempted to finish the experiment quickly). With these thoughts in mind, we defined the allowed time for responding on each trial in the Short condition as the average of the three lowest response times observed (this average was 1119ms) and in the Long condition as the average of the three highest average times observed (5096ms).

Regarding the pilot results, two brief notes are useful. First, only two participants played through the entire set of trials, without stopping. As we will shortly argue, meaningful results are only from participants who stopped before the highest possible trial, 75. Thus, it is encouraging regarding the robustness of the task that this was the case for most participants. Second, for participants who did stop before trial 75, the average stated probability that the next trial was *Y* was 74.3% (which was significantly different from 50%, $t(8)=4.71$, $p=.002$). Again, this is consistent with general expectation for a properly working experimental task.

For the remaining 28 participants, the instructions stated that there would be a fixed amount of time to respond on each trial and that participants should use all the available time to decide. They were also told that they would be prompted on the screen for when to respond. The trials were the same as for the pilot participants, but for the following difference: participants could not respond until they were prompted to do so (with the sentence 'please decide now', appearing in the middle of the computer screen). Participants were meant to respond as soon as possible, after the prompt had appeared, but we could not restrict the allowed response time (since we could not automatically advance the game to the next trial). The prompt 'please decide now' appeared 1119ms after the beginning of the trial in the Short condition and 5096ms in the Long condition.

Finally, participants were asked to complete Carver and White's (1994) BIS/BAS questionnaire, which is comprised of four scales: a behavioral inhibition system, relating to aversive goals, and a Behavioral Approach System, composed drive, fun seeking, and reward responsiveness scales. The BIS/BAS questionnaire was included as an exploratory tool, to consider individual differences, which might moderate performance in the doubling game task. For example, perhaps high scores on the BIS scale might associate with a lower stopping trial, while high scores on reward responsiveness a higher stopping trial.

3.3 Results

We seek to test the prediction of the QP formulation, that participants in the Long condition are likely to stop the game at a higher trial, compared to participants in the Short condition.

The experimental design assumes that participants utilize the allowed time per trial to consider whether they should continue playing or not and, equally, that whether they decide to continue to play or not will depend on a tradeoff between a desire for higher gains and a fear that the next trial would be *Y*. Regarding the latter, if the probability that *Y* is the next trial is too low, then it would make no sense to stop playing (even for loss averse participants).

Obviously, these are idealized assumptions. The collected data provide three ways in which the assumptions can be explored. First, we can look at the average response time, once the 'please decide now' prompt had appeared on the screen. We observed a range of average response times from 462ms to 5775ms, indicating that some participants simply ignored the instructions. Moreover, the average response time did depend on condition, so that participants in the Short condition required more time to respond (2191ms), than the ones in the Long condition (1081ms; $t(26)=2.93$, $p=.007$), suggesting that participants in the Short condition compensated for the shorter allowed time by taking more time to respond, after the 'please respond now' prompt (cf. Pleskac & Busemeyer's, 2010, finding of a positive relationship between confidence and decision time). One way to address this confound is to eliminate participants with a response time of more than 2000ms (this is a guide response time, for responding to the question, assuming a decision had already been reached). Second, participants who played the game beyond trial 75 might have done so for a variety of reasons, other than prerogatives assumed relevant to the problem. For example, they might be curious to see how long the experiment lasts or guess that a *Y* does not actually exist. Such explicit strategies confound the effect of interest. There were two such participants and we decided to eliminate their results from the main analyses. Third, we can consider the stated probability that, after a participant had stopped playing, the next trial is *Y*. Even for loss averse decision makers, we expect that stopping behavior should not occur, when this probability is too low.

Exploring the main hypothesis of interest first, the main dependent variable is the last played trial. An independent samples t-test including all participants, comparing the average last trial between the short (38.9) and long (32.6) conditions, did not reach significance, $t(26)=1.135$, $p=.27$. We next eliminated all participants whose pattern of responses was suspicious, as discussed above. We eliminated the data from participants with response times of 2000ms or more and stated probability that the next trial is *Y* of 10% or less. With this subset of the participants, the average last trial for participants in the Short condition (33) and the Long condition (32) converged, if anything, and the result was still not significant ($t(18)=0.196$, $p=.85$). In both cases, the trend was in a direction opposite to that predicted.

Finally, we considered interesting aspects of the data, irrespective of the main hypothesis, and these analyses were carried out on the entire experimental sample (28 participants). First, having decided to stop playing, the average stated probability that the next trial is *Y* was 49.3%, a value not significantly different from 50% ($t(25)=.16$, $p=.87$; this analysis applies only to the 26 participants, who stopped prior to trial 75). Thus, there is good indication that the collective behavior of participants was broadly consistent with a value maximization approach in the experiment. However, note that there was considerable variability in these probabilities too (SD=22.9, the range was between 1% and 90%). Second, we explore putative relations between the BIS/BAS variables and last played trial, stated probability that the next trial is *Y*, and response times (Table 3). Regarding reward responsiveness, there is a trend approaching significance for a negative association with average response time (perhaps indicating impatience to proceed to the next trial?). Also, participants higher on BIS would stop playing, while expecting the next trial to be *Y* with a higher probability, perhaps suggesting that participants were overestimating the chance of ruin, prior to stopping. We note that no correction was carried out for multiple tests, but the intention here is simply to provide some baseline individual differences results.

Table 3. Correlations between the BIS/BAS indices and various variables from the doubling game results.

		Last trial	Average response time	Prob that the next trial is Y
BAS drive	r	.200	-.162	.066
	N	27	27	25
BAS fun seeking	r	.123	-.230	-.119
	N	27	27	25
BAS reward responsiveness	r	.261	-.350*	.242
	N	27	27	25

BIS	r	-.079	-.119	0.573**
	N	27	27	25

Note. The single star indicates a p-value approaching significance (p=.074) and the two stars a significant p-value (p=.003).

4. Discussion

Problems of vagueness in decision making, exemplified by the simple doubling game we considered, can lead to profound normative problems. Such problems have long been recognized, but in more verbal and less quantitative forms. For example, the Sorites paradox refers to the problem of deciding whether there is a heap of sand or not. Suppose a person starts with enough sand, so she can say without doubt that she has a heap. Then, she removes only one grain of sand from the heap. Surely, one grain of sand will not make the difference between having a heap and not having a heap. However, applying this procedure/ reasoning indefinitely, means that eventually the person will end up with just a single remaining grain of sand, at which point she surely no longer has a heap. Thus, there is a conflict between a local perspective to the problem ('removing one grain of sand will not make a difference in whether we have a heap or not') and a global one ('eventually there will be one grain of sand, and so there would be no heap'). Equally, cycles of strict preferences present us with an analogous problem, as Gilboa's (2009) example shows. Suppose that, if someone has eaten i peanuts, she would prefer to eat i+1 peanuts. But, equally, she does not want to eat more than 100 peanuts. This is a situation in which the person will always choose to have an extra peanut, but eventually she will realize she has eaten too many. Gilboa (2009) describes this conflict as a competition between a short-run payoff and a long-term goal. In any case, the problems in decision making which arise from vagueness are in our view more clearly highlighted in quantitative examples, because some researchers have argued that the Sorites paradox and other similar paradoxes are due to the vagueness of language (e.g., Keefe & Smith, 1997, Chapter 1).

Our approach focused on the conflict between the local and global perspectives, or goals, or lines of reasoning. While we have focused on the doubling game, we think that these ideas are general and could be applied in, for example, the well-known Exchange game. Importantly, in life people are often forced to deal with problems of vagueness (e.g. should I marry K? take job J?). Presumably we have evolved to make decisions in the face of vague specifications. People are mostly able to decide whether there is a heap of sand or not or whether a person is tall or not (e.g., Alxatib & Pelletier, 2011). We typically can and must make decisions in the face of vagueness, even if all the conflicting perspectives are well understood. For example, a participant would understand the conflicting local, global perspectives in the doubling game. We have a strong intuition that he is still likely to stop at some moderately high trial, though possibly having no idea how that trial should be chosen, at the outset of the game.

We developed a QP theory model for the doubling game, based on this idea of conflicting intuitions between a bias to continue playing on each trial and a desire to stop eventually. The former is implemented as a transfer of amplitude from the current stopping rule to higher stopping rules on every trial. This means that on every trial confidence in the current stopping rule is undermined and in higher stopping rules increases. The model can produce stopping behavior, if the bias to continue playing has not sufficiently undermined the belief in the latest stopping rule. Clearly, this is not a normative solution (e.g. a 'best' N on which to stop playing) for the doubling and related games. Rather, we characterize the dynamics of a person facing such decisions because, paradoxes of reasoning notwithstanding, when a decision must be made, people do make a decision.

As far as we know, there have not been previous attempts to model behavior in the doubling game and so no prior guidance. As such, the present proposal is best seen as a 'methods' model, that is, a framework which provides us with hypotheses of what factors could be affecting performance in a doubling game task, whereas before we had no idea of what could be such factors. So, for example, the model led us to predict that, when the time per trial is short, the transfer of amplitude away from the current stopping rule to higher ones is modest. Thus, if a collapse occurs, decision makers are most likely to reaffirm their latest stopping rule, which implies fewer revisions of the stopping rule and a lower overall stopping trial. Such a 'methods' model (or perhaps a better characterization would be an exploratory model) seems consistent with the spirit of this special issue, insofar that the work of Estes has shown how we can use mathematical modeling to generate insights about psychological process, which have been impossible otherwise.

Similar considerations apply to the experimental illustration we provided. The translation of the abstract doubling game into a laboratory task was not straightforward. The main experimental hypothesis was not supported. One possibility is that the QP theory model, as currently specified, is not correct. One suspect assumption is that the probability of collapse depends on the number of trials; that is, the model tosses an appropriately weighted coin to decide whether a collapse should occur or not on every trial. An alternative assumption might be that the model should check whether a collapse occurs or not, per unit time. If this latter possibility is more accurate, the rationale for the manipulation in the experiment would not apply and alternative tests of the quantum model would need to be considered.

Equally, the present results also enabled the identification of empirical challenges, which need be clarified, before the case is settled one way or another. How can we ensure that information about Y (the trial number at which everything is lost) is vague? Without vagueness, a player can generate a Bayesian prior and use this to guide behavior. With human participants coming to the lab for a psychology experiment, there are baseline expectations of what Y would be anyway (e.g., Y cannot be so large that the experiment will last a whole day). Relatedly, there is an intuition that, as Y becomes (in principle) smaller, then so does vagueness disappear and a Bayesian prior is sensible. Moreover, how can we ensure that participants have an unbounded desire for high gains (and indeed understand the notion of expected value, which is necessary for the normative paradoxes to emerge)? In a role-playing version of the doubling game, in which participants are asked to imagine winning real sums of money, they may stop playing a doubling game, when they have accumulated gains which are large, compared to their material needs. Equally, in a laboratory task, participants may stop early simply to finish the experiment. Any manipulation to encourage serious engagement with the task will need to concern returns and not, for example, maximizing the probability of non-ruin (in the latter case, the problem changes). Setting aside such empirical challenges of a more conceptual nature, we encountered concrete problems too. Notably, participants in the Short condition consistently produced longer response times, than participants in the Long one, casting doubt on whether the intended manipulation of restricting the allowed time per trial worked at all (see also Pleskac & Busemeyer, 2010). Moreover, in the Long condition, informal debriefing after the study indicated that participants would often get somewhat impatient with the wait for each trial. Was this a factor which encouraged them to play fewer trials than they may have otherwise done? Of course, what constitutes a Long vs. Short trial is participant dependent and perhaps a more appropriate approach would be to tailor the corresponding response times based on individual participant pilot data. These empirical challenges cast doubt on any strong conclusion regarding the QP theory model.

Note, the QP theory model can generate additional predictions. For example, suppose that a player is briefly distracted (say her telephone rings) or she is motivated to think about her decision (e.g., she is told to think more carefully before deciding to continue playing on the next trial). In the QP theory model, a distraction would either slow down or freeze the evolution of the state vector. Less thorough deliberation could be modeled as a smaller change in the evolution of the state vector, between game trials.

Is it possible that a Markov (CP theory) model could be specified, that broadly mimics the QP theory model? In such a Markov model, a vector could represent all possible stopping rules, with each entry indicating the probability of each corresponding stopping rule. A classical transition matrix could also be identified, such that it decreases the probability of the first e.g. k states, while the probability of the remaining j states increases. We could next assume that a classical player acts on the basis of her latest guess for a stopping rule, while she is not aware of any change in probability in favor of higher stopping rules (we discuss this shortly). We can also require that at random time points, the player decides to review her stopping rule, by tossing a coin with odds set by the way the probability weight is distributed amongst the possible stopping rules). The player then either re-affirms her preference for her current stopping rule or decides to consider a higher stopping rule. Finally, we can assume that, right after such a decision, all of the player's confidence is concentrated on the chosen stopping rule.

So, the main ideas in the QP theory model can be implemented classically. In fact, we can also specify a Bayesian model, which can generate a prescription for stopping, based on backward induction (Appendix 2). An exception where the classical approach requires an extra assumption concerns the requirement that the player is unaware of the accumulating probability weight, in favor of higher stopping rules. For a CP theory model this is an assumption, which needs to be introduced over and above prescription from CP theory. In QP theory, when the state representing possible stopping rules is a superposition state, then a specific value for a stopping rule does not exist (this is the implication of the Kochen-Specker theorem, mentioned above). Thus, as the player deliberates the evidence against the current stopping rule and in favor of higher ones, as long as the state is a superposition one, there are no specific possibilities for a higher stopping rule to be aware of. This in turns makes plausible the idea that she would be unaware of the accumulating evidence for each possible stopping rule. Why is it important to assume that the player (classical or quantum) is unaware of the changing evidence in favor of higher stopping rules, with every trial? If the player is unable to estimate the accumulated bias against her current stopping rule, she can stop playing on the basis of her current stopping rule, even if considerable weight has accumulated in favor of *higher* stopping rules. By contrast, if he is assumed to be aware of changes in probability in favor of higher stopping rules, then plausibly the predicted behavior would be more rigid (making it likely that the player will either stop using the first stopping rule or not at all).

The QP model differs from a putative matched CP one in other respects too, notably in that in the former changes in amplitude oscillate with time, while this is not the case in the latter. In a QP model, given enough time, any changes in the amplitudes of the state vector will be reversed. If empirical results could be obtained which do reveal such an oscillatory nature of e.g. confidence in decisions, then these would provide very strong support for the QP theory model. However, this is an empirical challenge which we have yet to overcome. Moreover, there seem to be good reasons to consider the behavior of dynamic QP theory models up to the first turning point only (Pothos & Busemeyer, 2009).

Here, if the simplicity of the approach were the sole criterion in deciding between a CP theory model vs. a QP theory one, we admit that there would be only a minor advantage for the quantum model. But, QP theory has proved useful in characterizing decision making, which appears problematic under standard formalisms (notably CP theory). Thus, providing some groundwork for extending that approach to the arena of decision making under vagueness appears worthwhile (Aerts, 2009; Aerts & Gabora, 2005; Atmanspacher et al., 2004; Busemeyer et al., 2011; Pothos & Busemeyer, 2009; Trueblood & Busemeyer, 2012; Wang & Busemeyer, in press; see Busemeyer & Bruza, 2012, or Khrennikov, 2010, for overviews).

Can a normative solution for the doubling game paradox be derived from QP theory? In a QP theory model, there is not a single set of priors. Recall, incompatible questions are questions for which answering one with certainty implies uncertainty regarding the other. Thus, if we have two incompatible questions A and B, the (effective) sample space and priors from A's perspective can be quite different from those from B's perspective. Could a normative solution emerge by considering appropriately defined perspectives, such as local and global ones? The answer to this question involves deep issues concerning what is to be 'gained' after a decision trial, whether this is phrased as subjective expected utility, or something else. In QP theory, defining utility is not easy. One would need to define an observable for utility and a corresponding operator. Then, the expected value of this observable could be assessed, relative to a state vector. What would be an appropriate utility observable is a difficult issue, since many of the traditional constraints in classical approaches to utility (e.g., various forms of monotonicity or transitivity; Gilboa, 2009) are typically violated in quantum theory (cf. LaMura, 2009).

To conclude, we started from a well-known conceptual paradox in decision theory. We specified a preliminary mathematical and empirical framework for its study. Our results show several challenges which need be overcome, before compelling conclusions are possible. Nevertheless, the present work has allowed a range of novel insights, in relation to both cognitive process and experimental design, which we hope will encourage further work in this important area.

Acknowledgments

We would like to thank Tim Hollowood for his help with the Hamiltonian and some of the corresponding analytical solutions. EMP was supported by Leverhulme Trust grant RPG-2013-004 and JRB by NSF grant ECCS – 1002188. EMP and JRB were supported by Air Force Office of Scientific Research (AFOSR), Air Force Material Command, USAF, grants FA 8655-13-1-3044 and FA 9550-12-1-0397 respectively. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

References

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

Aerts, D., & Gabora, L. (2005). A theory of concepts and their combinations II: A Hilbert space representation. Kybernetes, 34, 192-221.

Alxatib, S., & Pelletier, J. (2011). On the psychology of truth-gaps. In R. Nouwen, R. van Rooij, U. Sauerland, & H.-C. Schmitz (Eds.), Vagueness in Communication (pp. 13-36). Berlin: Springer.

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

Atmanspacher, H., Romer, H., & Wallach, H. (2006). Weak quantum theory: formal framework and selected applications. Weak quantum theory: complementarity and entanglement in physics and beyond. Foundations of Physics, 32, 379-406.

Barkan, R. & Busemeyer, J. R. (2003) Modeling dynamic inconsistency with a changing reference point. Journal of Behavioral Decision Making, 16, 235-255.

Blutner, R., Pothos, E. M., & Bruza, P. (under review). A quantum probability perspective on borderline vagueness.

Bonini, N., Osherson, D., Viale, R., & Williamson, T. (1999). On the psychology of vague predicates. Mind & Language, 14, 377-393.

Bruza, P. D., Kitto, K., Nelson, D. & McEvoy, C. L. (2009). Is there something quantum-like about the human mental lexicon? Journal of Mathematical Psychology, vol. 53, pp. 362-377.

Busemeyer, J. R., Wang, Z., & Townsend, J. T. (2006). Quantum dynamics of human decision-making. Journal of Mathematical Psychology, 50, 220-241.

Busemeyer, J. R., Pothos, E. & Franco, R., Trueblood, J. S. (2011) A quantum theoretical explanation for probability judgment 'errors'. Psychological Review, 118 (2), 193-218.

Busemeyer, J. R. & Bruza, P. (2012). Quantum models of cognition and decision making. Cambridge University Press: Cambridge, UK.

Busemeyer, J. R., Weg, E., Barkan, R., Li, X., & Ma, Z. (2000). Dynamic and consequential consistency of choices between paths of decision trees. Journal of Experimental Psychology: General, 129, 530-545.

Carver, C.S., & White, T.L. (1994). Behavioral inhibition, behavioural activation, and affective responses to impending reward and punishment: The BIS/BAS scales. Journal of Personality and Social Psychology, 67, 319-333.

Chalmers, D. J. (2002). The St. Petersburg two-envelope paradox. Analysis, 62, 155–157.

Gilboa, I. (2009). Theory of decision under uncertainty. Cambridge University Press: Cambridge.

Hughes, R.I.G. (1989). The Structure and Interpretation of Quantum Mechanics. Cambridge, MA: Harvard University Press.

Hyde, D. (1997). From heaps and gaps to heaps of gluts. Mind, 106, 641.
Isham, C. J. (1989). Lectures on quantum theory. Singapore: World Scientific.

Keefe, R. & Smith, P. (1997). Vagueness: a reader. MIT Press: Boston.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. Linguistics and Philosophy, 30, 1-45.

Khrennikov, A. Y. (2010). Ubiquitous quantum structure: From psychology to finance. Springer.

LaMura, P. (2009) Projective expected utility. Journal of Mathematical Psychology, 53, 408-414.

Paulus, M. P. & Yu, A. J. (2012). Emotion and decision-making: affect-driven belief systems in anxiety and depression. Trends in Cognitive Sciences, 16, 476-483.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. Psychological Review, 117, 864-901.

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.

Sharot, T., Velasquez, C.M., & Dolan, R.J. (2010). Do decisions shape preference? : Evidence from blind choice, Psychological Science, 21(9) 1231–1235.

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

Wang, Z., & Busemeyer, J. R. (in press). A quantum question order model supported by empirical tests of an a priori and precise prediction. Topics in Cognitive Science.

Appendix 1. Analytical specification of the model.

The key technical component of the QP theory model for the doubling game is the unitary matrix, which determines how the state vector changes with time. This change with time is assumed to correspond to the thought process of a player, as he goes through trials in the doubling game task. The unitary matrix is computed from Schrödinger's equation, $i \cdot \frac{d|\psi\rangle}{dt} = H \cdot |\psi\rangle$, in which H is a linear operator, called the Hamiltonian. The Schrödinger equation corresponds to the Kolmogorov forward equation in classical dynamical models. All the information about the dynamical properties of a quantum system is contained within H (which is analogous to a classical transition matrix). The solution to Schrödinger's equation is $|\psi_t\rangle = e^{-i \cdot t \cdot H} \cdot |\psi_{tstart}\rangle = U(t) \cdot |\psi_{tstart}\rangle$. The operator $U(t)$ is another linear (and unitary) operator, and is obviously related to the Hamiltonian via $U(t) = e^{-i \cdot t \cdot H}$. Regarding a doubling game, $|\psi_t\rangle = U(tstep \cdot trials) \cdot |\psi_{tstart}\rangle = (\prod_{trials} U(tstep)) \cdot |\psi_{tstart}\rangle$, that is, the effect of each trial on the knowledge regarding the next guess for a stopping rule can be modeled by successive applications of $U(tstep)$ on the state vector corresponding to the first time point. So, analytically, the model specification requires us to derive a suitable Hamiltonian.

We first consider the specification of an appropriate H (and U) for the initial, wild guess. With this in place, we can derive the corresponding expressions for subsequent guesses. Let us then consider a state vector right after the initial guess of the form $|v\rangle^T = |1\rangle + |2\rangle + |3\rangle \ldots |initial\ guess\rangle + \cdots |Max\rangle$, where the number of states up to and including $|initial\ guess\rangle$ is p and the number of states higher than $|initial\ guess\rangle$ is q; obviously, $p+q=Max$ (note that for notational convenience we are still ignoring the necessary normalization, which would weigh each state by $1/\sqrt{Max}$; the superscript T denotes the transpose of a vector). We wish to specify a Hamiltonian such that the amplitude of the first p states will (initially) decrease and the amplitude of the remaining q states increase. Consider an H such that its first p rows have p 0's followed by q -1's, and the remaining q rows have p +1's, followed by q 0's, and all entries are multiplied by i. For example, for p=2, q=4, H would be i times:

$$\begin{matrix} 0 & 0 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{matrix}$$

In order to be able to use this Hamiltonian (this family of Hamiltonians) to derive a corresponding unitary matrix, we need to identify the so-called eigenvalues and eigenvectors of the Hamiltonian. In general, for any linear operator A, there may be a set of vectors such that operating A on these vectors transforms them onto a multiple of themselves, that is, $A \cdot x = \lambda \cdot x$. The vectors for which this occurs are called eigenvectors of A and the values λ are called eigenvalues. For the present purposes, we need to identify the eigenvalues, eigenvectors of the Hamiltonian in a way that they, ideally, depend only on p and q. Note that the eigenvalues and eigenvectors have no psychological meaning (in this case), they are just needed for the technical specification of the model. We first express the eigenvectors in a way which captures the structure of the Hamiltonian, as $v = (a\ a\ a\ a \ldots b\ b\ b\ b)^T$, where there would be a p number of a's and a q number of b's. Then, applying the usual rules of matrix multiplication, the eigenvalue equation is:

$$Hv = i\begin{pmatrix} -qb \\ \cdots \\ -qb \\ pa \\ \cdots \\ pa \end{pmatrix} = \lambda v$$

where λ are the eigenvalues. Thus, the final form of the eigenvalue equation is equivalent to $\begin{pmatrix} x \\ y \\ \cdots \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ \cdots \end{pmatrix}$ and, of course, we expect $x = x'$ etc. This leads to the equations $-iqb = \lambda a$ and $ipa = \lambda b$, which allow us to identify the two eigenvalues, $\pm\sqrt{pq} = \lambda$. Let us call v_+, v_- the eigenvectors corresponding to the positive and negative eigenvalues, respectively. Then, the corresponding eigenvectors are:

$$v_+ = \begin{pmatrix} \frac{1}{\sqrt{p}} \\ \cdots \\ \frac{1}{\sqrt{p}} \\ \frac{i}{\sqrt{q}} \\ \cdots \\ \frac{i}{\sqrt{q}} \end{pmatrix}, v_- = \begin{pmatrix} \frac{1}{\sqrt{p}} \\ \cdots \\ \frac{1}{\sqrt{p}} \\ \frac{-i}{\sqrt{q}} \\ \cdots \\ \frac{-i}{\sqrt{q}} \end{pmatrix}$$

Note that in both cases there are p terms of the form $\frac{1}{\sqrt{p}}$ and q terms of the form $\frac{i}{\sqrt{q}}$ or $\frac{-i}{\sqrt{q}}$. Recall, at the start of the game, the state vector can be written as $|\psi_{tstart}\rangle = \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix}$, ignoring normalization to simplify the subsequent computations. Note that each element of the state vector is the amplitude to a corresponding basis vector. Thus, the first element of the state vector is the amplitude for the $|1\rangle$ basis vector, the second element for the $|2\rangle$ basis vector etc. The simple form of the eigenvectors means that we can express the state vector as a combination of these eigenvectors. Specifically, we can write $|\psi_{tstart}\rangle = \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} = c_1 v_+ + c_2 v_-$, where, $c_1 = \frac{1}{2}(\sqrt{p} - i\sqrt{q})$, $c_2 = \frac{1}{2}(\sqrt{p} + i\sqrt{q})$. We have managed to express $|\psi_{tstart}\rangle$ in terms of the eigenvectors of the relevant Hamiltonian. This is extremely desirable in QP theory, since it allows us to compute $U(t) = e^{-i \cdot t \cdot H}$, in a simple way, as a function of the eigenvalues of H (otherwise, computing $U(t)$ from H requires a matrix exponentiation function).

Specifically, we have that (note that this equation is difficult to comprehend without a background in quantum theory):

$$|\psi_t\rangle = e^{-i \cdot t \cdot H}|\psi_{tstart}\rangle = \sum_k e^{-i \cdot t \cdot h_k}|h_k\rangle\langle h_k|\,|\psi_{tstart}\rangle = \sum_{k=+,-} e^{-i \cdot t \cdot v_k}|v_k\rangle\langle v_k|\,(c_1 v_+ + c_2 v_-)$$

where the summation ranges across all possible eigenvalues, in this case just the two $\lambda = \pm\sqrt{pq}$. This leads to $|\psi_t\rangle = c_1 e^{-i \cdot t \cdot \lambda_+} v_+ + c_2 e^{-i \cdot t \cdot \lambda_-} v_-$. We next use the identity $e^{i\theta} = cos\theta + isin\theta$ and, ignoring factors which are common to all the states, we have:

$$|\psi_t\rangle = (\sqrt{p} - i\sqrt{q})\{\cos(t \cdot \sqrt{pq}) - isin(t \cdot \sqrt{pq})\}v_+ + (\sqrt{p} + i\sqrt{q})\{\cos(t \cdot \sqrt{pq}) + isin(t \cdot \sqrt{pq})\}v_-$$

Here, we have an equation of the form $\begin{pmatrix} x \\ y \\ \dots \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ \dots \end{pmatrix} + \begin{pmatrix} x'' \\ y'' \\ \dots \end{pmatrix}$, which can be broken down into separate equations of the form $x = x' + x''$ etc. We can examine this equation separately for the first p terms of v_+ and v_- (which are, in both cases, $\frac{1}{\sqrt{p}}$) and the remaining q terms of v_+ and v_- (which are $\frac{i}{\sqrt{q}}$ and $\frac{-i}{\sqrt{q}}$, respectively). Recall, the amplitude for the first p terms is meant to be decreasing and the amplitude for the remaining q terms increasing. Indeed, after the necessary algebra, we find that the amplitude of the first p terms of $|\psi_t\rangle$ varies as $2\cos\left(t \cdot \sqrt{pq}\right) - 2\frac{\sqrt{q}}{\sqrt{p}}\, sin\left(t \cdot \sqrt{pq}\right)$ and the amplitude of the remaining q terms as $2\cos\left(t \cdot \sqrt{pq}\right) + 2\frac{\sqrt{p}}{\sqrt{q}} sin\left(t \cdot \sqrt{pq}\right)$. This completes the specification of the analytic solution for how the state vector varies with time. That is, we now know analytically how the amplitude for each basis vector varies with time.

We can use the above equations to compute the time at which the first turning point occurs (this is the same for both the increasing and decreasing amplitudes). Specifically, we computed the modulus of amplitude squared (that is, probability) and considered $\frac{d}{dt}(|amplitude|^2) = 0$. This produced the solutions

$$t = \frac{1}{2\sqrt{pq}}\left(tan^{-1}\left(\frac{2\sqrt{pq}}{q-p}\right) \pm \pi\right)$$

Note that the tangent function is periodic in π and whether we need to add π, subtract π, or neither, in order to identify the time for the first turning point, will depend on the size of p relative to q.

The computations so far correspond to what happens for the first guess, in which we assume that the state vector is maximally uniformed (the amplitude for all basis vectors is the same). This scheme can be readily extended to later guesses (in fact, many of the computations are simpler). After the latest guess, where that guess is not the first one, the relevant state vector would be $|\psi_{tstart}\rangle = \begin{pmatrix} 0 \\ 0 \\ \dots \\ guess \\ 0 \\ 0 \end{pmatrix}$, so that all amplitude/ probability weight is concentrated on the state corresponding to this latest guess. In order to compute the changes to the state vector, as a result of continuing wins, we again need to apply a suitable operator $U(t)$. However, we assume that once the player has committed herself to a particular guess for a stopping rule, she no longer considers the possibility of adopting a lower stopping rule. Therefore, the full space of possible stopping rules of dimensionality *Max* is no longer relevant, rather we are only interested in the subspace of stopping rules higher than the current guess. This can be formally implemented by expressing $|\psi_{tstart}\rangle$ as a direct sum decomposition of the form $\begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix} \oplus \begin{pmatrix} guess \\ \dots \\ 0 \end{pmatrix}$ and, likewise, writing the Hamiltonian as a direct sum of its restriction to the corresponding subspaces, so that $H = H_0 \oplus H_{guess}$, where H_0 is just a matrix of zeroes. This means that the part of $|\psi_{tstart}\rangle$ with basis vectors with indices lower than guess stays constant (at 0), while the rest of $|\psi_{tstart}\rangle$ evolves as above, that is, in the way specified for the player's, initial guess.

We first consider the relevant computations assuming that H_{guess} is the entire space, but with p=1, q=n-p, and later show how the results can be easily adapted to take into account the fact that H_{guess} is a subspace of *H*. In the computations below, *n* denotes the dimensionality of H_{guess}. Using the equations derived above, for the positive and negative eigenvalue of H_{guess}, we have respectively:

$$\lambda = \lambda_+ = \sqrt{n-1}, v_+ = \begin{pmatrix} 1 \\ \frac{i}{\sqrt{n-1}} \\ \frac{i}{\sqrt{n-1}} \\ \cdots \\ \frac{i}{\sqrt{n-1}} \end{pmatrix} \text{ and } \lambda = \lambda_- = -\sqrt{n-1}, v_- = \begin{pmatrix} 1 \\ \frac{-i}{\sqrt{n-1}} \\ \frac{-i}{\sqrt{n-1}} \\ \cdots \\ \frac{-i}{\sqrt{n-1}} \end{pmatrix}$$

As before, we seek to express $|\psi_{tstart}\rangle$, the state right after the (latest) collapse of the state vector, in terms of the eigenvectors of H_{guess}. We have here: $|\psi_{tstart}\rangle = \begin{pmatrix} 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix} = c_1' v_+ + c_2' v_-$, so that $c_1' = \frac{1}{2} = c_2'$. Therefore,

$$|\psi_t\rangle = \frac{1}{2} e^{-i \cdot t \cdot \lambda_+} v_+ + \frac{1}{2} e^{-i \cdot t \cdot \lambda_-} v_- \Longrightarrow$$

$$|\psi_t\rangle = \frac{1}{2}\{\cos\left(t \cdot \sqrt{n-1}\right) - isin\left(t \cdot \sqrt{n-1}\right)\} v_+ + \frac{1}{2}\{\cos\left(t \cdot \sqrt{n-1}\right) + isin\left(t \cdot \sqrt{n-1}\right)\} v_-$$

Repeating the analysis we carried out above for the first, wild guess, we find that the amplitude for the first term varies with time as $\cos\left(t \cdot \sqrt{n-1}\right)$. Note that the cosine function is initially decreasing. Moreover, the amplitude for the last n-1 terms varies as $\frac{1}{\sqrt{n-1}} sin\left(t \cdot \sqrt{n-1}\right)$. The sine function is initially increasing. It can be seen that the amplitude lost by the first state is basically shared by the remaining n-1 states, in a way that depends on the number of these remaining states (note that these computations include all appropriate normalization as well). Also, the first turning point clearly occurs when $t \cdot \sqrt{n-1} = \frac{\pi}{2} \Leftrightarrow t = \frac{\pi}{2\sqrt{n-1}}$. These expressions were derived by considering the reduced dimensionality subspace, H_{guess}, which is n. We next reformulate these expressions in terms of the dimensionality of the full space, H. This can be easily achieved by introducing the variable *guess*, which denotes the trial corresponding to the player's latest stopping rule. Then, the dimensionality of H_{guess} would be *Max*-*guess*+1, so that the first turning point would occur at $t = \frac{\pi}{2\sqrt{Max-guess}}$.

This completes the analytical specification of the model. We conclude this section by summarizing some algorithmic details of the model. Initially, the model picks a random stopping rule for the first, wild guess, and sets the state vector $|\psi\rangle$ to a maximally uninformed superposition. The *tmax* is computed and, given an externally set value of $ExCo$ as well, we can compute the probability for collapsing $|\psi\rangle$ on each trial. On each trial, the model first examines whether the trial count has reached the stopping rule. If yes, then model stops playing. If not, the model examines whether a collapse of $|\psi\rangle$ occurs on that trial or not, by flipping a coin, whose odds of success are set depending on the $ExCo$ parameter and *tmax*. That is, we assume that the probability of collapse depends on the number of trials, but not on the time per trial (after all, it is with each trial that the model acquires new information from the environment). If a collapse does not occur, the model proceeds to the next trial, changing the non-zero amplitudes of the state vector, as specified above. If a collapse does occur, then the model determines the next stopping rule (which could be the same as the existing one), by computing the probability for each eligible stopping rule, that is, all stopping rules with non-zero amplitude. In practice, we also need to ignore stopping rules whose probability is zero to a certain number of decimal places (four decimal places in the current simulations). Based on the probability for each stopping rule, the model randomly selects the next stopping rule. The state vector $|\psi\rangle$ is then set accordingly, so that all the amplitude is concentrated on the state corresponding to the selected stopping rule. Also, a new *tmax* is computed and so a new probability for collapsing the state vector on each trial. The model iterates until either the model stops (that is, the current trial counter reaches the latest stopping rule) or the maximum possible trial is reached (in the current simulations this was set to 1000).

Appendix 2. A simple Bayesian model to generate a prescription for a stopping trial, in the doubling game.

Let *x* be the stake, which is doubled after each play.

Suppose the player's prior probability distribution over the possible values of n = 1,2,.., N that the game stops and all is lost equals

$$P_0 = \begin{bmatrix} p_1 & \dots & p_k & \dots & p_N \end{bmatrix} \quad (1)$$

P_0 is the prior before any actual plays. p_k equals the probability that the game stops on play number *k.* N is the maximum possible trial the game can end. The game will end for sure on N, but it might end before N too.

After one play, if the game does not end, the prior is updated to the following posterior

$$P_1 = \begin{bmatrix} 0 & p_2 & p_3 & \dots & p_N \end{bmatrix} / S_1$$

$$S_1 = p_2 + \dots + p_N$$

After k plays, if the game does not end, the posterior equals

$$P_k = \begin{bmatrix} 0 & \dots & 0 & p_{k+1} & \dots & p_N \end{bmatrix} / S_k$$

$$S_k = \left(p_{k+1} + \dots + p_N \right)$$

(read this as drop the first *k* terms when computing *S*)

Now we do backward induction. This means that the player tries to figure out the optimal choice and the optimal payoff, given that choice, for every possible state of the world (not assuming any one is true, but using the probability of each possible state of the world, to determine the expectation; maximizing the expectation determines the optimal choice).

Stage N-1

The expected value to play again after N-1 plays without the game having ended equals $E_{N\ 1} = 0$, because the player believes the next play, at the final N, will mean everything is lost. So the optimal strategy is to stop at N-1 and win $E^*_{N\ 1} = 2^{N\ 1} \times x$ (rather than lose all for sure).

Stage N-2

The expected value to play again after N-2 plays without the game having ended equals

$$E_{N\ 2} = \left(p_N / S_{N\ 2} \right) \times E^*_{N\ 1} \quad (2)$$

This involves the probability that the game will end on trial N, and not on trial N-1, given that the game has not ended across trials 1 to N-2, times the gain, which would be realized after playing N-1 trials.
Note that $p_N = S_{N\ 1}$, so the above can be written as $E_{N\ 2} = \left(S_{N\ 1} / S_{N\ 2}\right) \times E*_{N\ 1}$.

If $E_{N\ 2} = \left(p_N / S_{N\ 2}\right) \cdot 2^{N\ 1} \cdot x > 2^{N\ 2} \cdot x \rightarrow \left(p_N / S_{N\ 2}\right) > 1/2$, then the player should be planning to play after N-2. That is, the optimal strategy is to continue at N-2 with $E*_{N\ 2} = E_{N\ 2}$ (i.e. Eq. 2), as the future expected payoff.

On the one hand, if $\left(p_N / S_{N\ 2}\right) < 1/2$ then the optimal strategy is to stop at N-2 and realize $E*_{N\ 2} = 2^{N\ 2} \times x$, as the future optimal payoff. This is not the same as Eq. 2.

<u>Stage N-3</u>
The expected value to play after N-3 plays, assuming the game has not ended up to that point, equals

$$E_{N\ 3} = \left(S_{N\ 2} / S_{N\ 3}\right) \times E*_{N\ 2}$$

$S_{N\ 2} = p_{N\ 1} + p_N$. That is, for the player to want to play after N-3, he needs to assume that the game will stop either on trial N-1 or N, not on trial N-2.

$S_{N\ 3} = p_{N\ 2} + p_{N\ 1} + p_N$

The probability that the game ends after trial N-3 is $p_{N\ 2}$. The probability that it does not end on trial N-3 (given that the game has gotten this far) equals S_{N-2} (shown above). The probability of stopping anywhere (given that the game has reached trial N-3) is S_{N-3} (shown above).

If $E*_{N\ 3} = \left(S_{N\ 2} / S_{N\ 3}\right) \times E*_{N\ 2} > 2^{N\ 3} \times x$, then the player plays after N-3. Therefore, the player plays again if $\left(S_{N\ 2} / S_{N\ 3}\right) > \left(2^{N\ 3} \times x / E*_{N\ 2}\right)$.

Note, $S_{N-2}/S_{N-3} \cdot E^*_{N-2}$ is the expected gain if the player continues to play from trial N-3 and $2^{N-3} \cdot x$ is the guaranteed again, assuming the player has reached N-3 and decides to stop there.

<u>Stage N-j</u>

The expected value to play after N-j plays, assuming the game has not ended up to that point, equals

$$E_{N\ j} = \left(S_{N\ j+1} / S_{N\ j}\right) \times E*_{N\ j+1}$$

If $E*_{N-j} = \left(S_{N-j+1} / S_{N-j}\right) \times E*_{N-j+1} > 2^{N-j} \times x$, then the player plays after N-j. Therefore, the player plays again if $\left(S_{N-j+1} / S_{N-j}\right) > \left(2^{N-j} \times x / E*_{N-j+1}\right)$.

The above model is based on backward induction, which is about how a player makes a plan for the game, at the outset (this approach is like dynamic programming, a standard method for solving for optimal decisions that involve a sequence of actions and events across trials). In order to figure out what to do on the first trial, the player starts from the last trial, assuming he has won all previous trials, and continues backwards, until he reaches his current first trial. If he decides to play the next trial and wins, then the plan put together initially already prescribes an action for whether to play on the third trial or not, since the initial plan for trial 3 was specified assuming a win on trial 2 (that is, the plan for deciding whether to continue playing from trial 2 or not was already based on posterior probabilities, assuming winning on trial 1).

Such a Bayesian model seems most applicable in situations whereby a player is introduced to the doubling game and, prior to actually starting to play the doubling game, is asked to decide on a stopping rule. How well would we expect such a simple Bayesian model to work? Suppose the player decides he should stop playing at trial k. In actually reaching trial k-1 and not having lost, we suggest that the player would plausibly no longer believe in the original prior and, so, that he should not stop at trial k-1. Why? Because having actually won k-1 trials, it is plausible that the player will expect he can continue winning. If these intuitions are correct, then such a Bayesian model is unlikely to be a good model of human behavior, because there is no mechanism for changing the prior; perhaps players in such situations do not decide on a prior in advance at all (see also Busemeyer et al., 2000).

Notwithstanding these qualifications, we wanted to illustrate that a prescription for a stopping trial does not require a quantum approach and can be generated from a simple Bayesian model.