Invariants in probabilistic reasoning

Fintan Costello[a],[*], Paul Watts[b]

[a] *School of Computer Science and Informatics, University College Dublin, Belfield, Dublin 4, Ireland*
[b] *Department of Theoretical Physics, National University of Ireland Maynooth, Maynooth, Co Kildare, Ireland*

A R T I C L E   I N F O

Keywords:
Rationality
Biases
Probability

A B S T R A C T

Recent research has identified three invariants or identities that appear to hold in people's probabilistic reasoning: the QQ identity, the addition law identity, and the Bayes rule identity (Costello and Watts, 2014, 2016a, Fisher and Wolfe, 2014, Wang and Busemeyer, 2013, Wang et al., 2014). Each of these identities represent specific agreement with the requirements of normative probability theory; strikingly, these identities seem to hold in people's judgements despite the presence of strong and systematic biases against the requirements of normative probability theory in those very same judgements. These results suggest that the systematic biases seen in people's probabilistic reasoning follow mathematical rules: for these particular identities, these rules cause an overall cancellation of biases and so produce agreement with normative requirements. We assess two competing mathematical models of probabilistic reasoning (the 'probability theory plus noise' model and the 'quantum probability' model) in terms of their ability to account for this pattern of systematic biases and invariant identities.

1. Introduction

A fundamental goal of science is to find invariants: constant mathematical relationships that hold between different variables (Simon, 1990). Such invariants are perhaps the defining characteristic of quantitative 'hard' science: almost all important results in physics (Maxwell's equations of electromagnetism, Newton's law of gravitation, Einstein's field equations, the mass-energy equivalence, and so on) describe such invariant relationships. While such mathematical invariants or identities are rarer in the behavioural sciences, recent work has identified three identities that appear to hold in people's intuitive probabilistic reasoning: the 'QQ' ('Quantum Question') identity (Wang & Busemeyer, 2013; Wang, Solloway, Shiffrin, & Busemeyer, 2014), the addition law identity (Costello & Watts, 2014; Fisher & Wolfe, 2014), and the Bayes rule identity (Fisher & Wolfe, 2014; Costello & Watts, 2016a). Each identity describes a constant relationship that holds between different probabilistic judgements, and each represents specific agreement with the requirements of classical probability theory in those judgements. Strikingly, these identities or invariants hold in people's intuitive judgements of probability despite the presence of strong biases, or systematic deviations from the requirements of probability theory, in those very same judgements.

The fact that these mathematical identities appear to hold in people's probabilistic judgement (alongside patterns of systematic bias in those same judgements) has important implications for our understanding of how people reason about probability. It suggests that people judge probability in a way that follows some sort of formal, mathematical process that causes systematic biases in

---

judgement whose values cancel each other out in particular invariant relationships. Indeed, these patterns of systematic bias alongside invariant identities are predicted by two competing formal models of probabilistic reasoning. One model, based on noisy frequentist probability, predicts the addition law and Bayes rule identities (Costello & Watts, 2014, 2016a) while the other, based on quantum probability, predicts the QQ identity (Wang & Busemeyer, 2013; Wang et al., 2014). In this paper we ask whether either of these models are able to account for all three invariant identities. In the first section we explain the QQ, addition law and Bayes rule identities. In the second section we present the quantum probability model and show that, while this model predicts the QQ identity, it is fundamentally unable to account for addition law and Bayes rule identity results. In the third section we present the 'probability theory plus noise' model and explain how it predicts the addition law and Bayes rule identity results. We also describe how priming effects in this model allow it to explain the QQ identity and to make some novel predictions connected to that identity (predictions that are supported by experimental results). In the fourth section we draw some general conclusions.

2. Identities in probabilistic reasoning

In presenting the addition law, Bayes rule and QQ identities we use the following notation, derived in part from quantum probability theory. We take $P(A)$ to represent the normatively correct probability of event A. We take **A** to represent a question about the occurrence or non-occurrence of A: in the language of quantum probability, **A** is an 'observable' that returns either A or $\neg A$. The QQ identity is an invariant that relates answers to questions presented in different sequential orders. For tasks involving question ordering, we consider two possible orderings represented as **AB** (first a question about A, then a question about B) or **BA** (first a question about B, then a question about A). We take $P_{\mathbf{AB}}(A)$ to represent the subjective *estimated* probability of A in the first ordering (the probability of getting a 'yes' answer to the question **A** when **A** is asked first and **B** is asked second). We take $P_{\mathbf{BA}}(A)$ to represent the subjective estimated probability of A in the second ordering (the probability of getting a 'yes' answer to the question **A** when **B** is asked first and **A** second). We take $P_*(A)$ to represent the subjective estimated probability of A when order is irrelevant (when **A** and **B** do not occur sequentially).[1] Since a subsequent presentation of question **B** cannot affect the results obtained from a prior presentation of question **A** (time travel is not allowed!), $P_*(A) = P_{\mathbf{AB}}(A)$ and $P_*(B) = P_{\mathbf{BA}}(B)$.

2.1. The QQ identity

Consider a situation where people are asked yes-no questions in two alternative orders **AB** or **BA**. This situation is commonly seen in polls; for example, in a Gallup poll conducted in September 1997, half of participants were asked the question "Do you think Al Gore is honest and trustworthy?" followed immediately by the question "Do you think Bill Clinton is honest and trustworthy?", while the other half of participants were asked the same questions in the reverse order (Moore, 2002). A noticeable pattern of bias in such situations is that people's answers for a given question are often strongly influenced by the order of question presentation: the probability of a 'yes' answer to question **A** when that question comes first can be significantly different from the probability of a 'yes' answer when question **A** is preceded by question **B** ($P_{\mathbf{AB}}(A) \neq P_{\mathbf{BA}}(A)$). In the Clinton-Gore questions, for example, 76% of participants answered 'yes' to the Gore question when it was asked first (the **AB** order; $P_{\mathbf{AB}}(A) = 0.76$), while 66% answered yes when that question was asked second, after the Clinton question (the **BA** order; $P_{\mathbf{BA}}(A) = 0.66$): the prior presentation of the Clinton question produced a bias, reducing the likelihood of a 'yes' answer to the Gore question. These order effects occur both in polls on a range of different topics (Moore, 2002) and in similar experimental studies (Wang & Busemeyer, 2013; Wang et al., 2014).

Simultaneously, however, results (both from experimental studies and from poll data) show that the following identity tends to hold reliably in such sequential question answering:

$$P_{\mathbf{AB}}(A \wedge B) + P_{\mathbf{AB}}(\neg A \wedge \neg B) - P_{\mathbf{BA}}(A \wedge B) - P_{\mathbf{BA}}(\neg A \wedge \neg B) = 0$$

(this expression has a value of -0.003 in answers to the Clinton-Gore questions, for example). This represents unbiased agreement with the requirements of probability theory (in which, of course, the probability of a conjunction $A \wedge B$ does not depend on the order of the events within the conjunction). This identity appears to hold for all such consecutive questions, despite significant order effects for the same set of question answers. This identity holds for questions across a wide range of different topics in 72 different national representative surveys in the US, and in laboratory studies of the effects of order in question answering (Wang et al., 2014). This identity does not appear to hold when questions are not consecutive.

To connect the QQ identity more closely to order effects in sequential judgement, we rewrite it in the form

$$P_{\mathbf{AB}}(A \wedge B) - P_{\mathbf{BA}}(A \wedge B) = -[P_{\mathbf{AB}}(\neg A \wedge \neg B) - P_{\mathbf{BA}}(\neg A \wedge \neg B)] \tag{1}$$

This identity reveals an interesting pattern in the effects caused by question ordering: the contextual order effect observed when taking the difference between the probability of answering 'yes' to both questions in the order **AB** and the probability of answering 'yes' to both questions the order **BA** is equal to the negative of the contextual order effect observed when taking the difference between the probability of answering 'no' to both questions in the order **AB** and the probability of answering 'no' to both questions
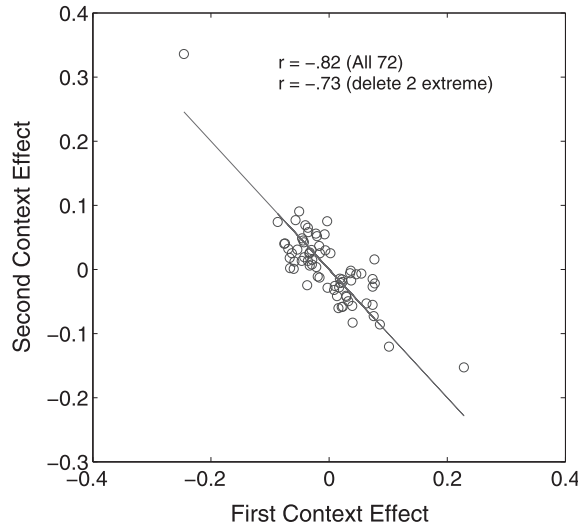

Fig. 1. Scatterplot of contextual order effects for the 72 studies in Wang et al. (2014). Each point represents one of these studies, with the *x*-axis giving the difference between the probability of answering 'yes' to both questions **A**,**B** in the order **AB** and the probability of answering 'yes' to both questions in the order **BA**. This first contextual order effect is equal to the probability difference $P_{AB}(A \wedge B) - P_{BA}(A \wedge B)$ for that study. The *y*-axis gives the difference between the probability of answering 'no' to both questions in the order **AB** versus the probability of answering 'no' to both questions in the order **BA**. This second contextual order effect is equal to the probability difference $P_{AB}(\neg A \wedge \neg B) - P_{BA}(\neg A \wedge \neg B)$ for that study. The quantum probability model predicts that these points will fall along a line with intercept 0 and slope −1 (solid line in the Figure). This image is taken from Wang et al. (2014).

the order **BA**. A scatterplot of values of the probability differences on the left and right sides of Eq. (1), for the 72 studies analysed by Wang et al. (2014), is shown in Fig. 1: as the scatterplot shows, these differences are related as in Eq. (1).

While this QQ identity is stated in terms of the probabilities of binary choices (the probability of answering 'yes' or 'no' to a given question), recent work has tested a very similar QQ identity in probability estimation (Yearsley & Trueblood, 2017). In a study examining people's judgements of the probability of various candidates winning the Democratic or Republican nomination (*N*) for president in the 2016 US elections given that they had won or lost two primaries *A* and *B*, Yearsley and Trueblood (2017) found a reliable effect of order of presentation such that $P(N|A,B) \neq P(N|B,A)$ (that is, such that the probability of a candidate winning the nomination given that they had won primary *A* and primary *B*, presented in that order, was reliably different from the probability of the candidate winning the nomination given that they had won primary *B* and primary *A* presented in that order). Despite this order effect, however, Yearsley and Trueblood (2017) also found that the following version of the QQ identity

$$P_{AB}(N|A \wedge B) + P_{AB}(N| \neg A \wedge \neg B) - P_{BA}(N|A \wedge B) - P_{BA}(N| \neg A \wedge \neg B) = 0$$

held in probability estimates, on average (where, for example, $P_{AB}(N|A \wedge B)$ represents the probability of *N* given that *A* and *B* hold and they are presented in the order **AB**).

This general pattern, of bias due to order of question presentation ($P_{AB}(A) \neq P_{BA}(A)$) combined with simultaneous agreement with the QQ identity, is predicted by a model of reasoning based on quantum probability (Wang & Busemeyer, 2013; Wang et al., 2014). The fact that this QQ identity appears to hold for consecutive questions, despite significant order effects for the same set of questions, has been seen as providing 'the strongest form of support [for the quantum probability model] because its prediction is not dependent on parameter choices-other applications of quantum theory to human cognitive depend on choosing parameter values to best fit data' (Wang et al., 2014).

2.2. The addition law identity

Probability theory requires that certain identities must hold for probability estimates involving any pair of events *A* and *B*. One such identity is the addition law, which requires that

$$P(A) + P(B) - P(A \wedge B) - P(A \vee B) = 0 \tag{2}$$

must hold for all events *A* and *B*. Two other 'expansion' identities require that

$$P(A \wedge B) + P(A \wedge \neg B) - P(A) = 0 \tag{3}$$

and

$$P(A \wedge B) + P(\neg A \wedge B) - P(B) = 0 \tag{4}$$

must hold for all events *A* and *B*. The addition law identity follows from the rules of set intersection and union represented by the $A \wedge B$ and $A \vee B$ terms; the two expansion identities follow from the fact that a given set of events *A* consists of two mutually

Table 1

Predicted values of the noise model and the quantum model for a series of probability theory identities. Standard probability theory requires these identities to have a value of 0. In applying the quantum probability model to these identities, $P(A)$ and $P(B)$ are always measured in orderings $P_{\mathbf{AB}}(A)$ and $P_{\mathbf{BA}}(B)$ respectively. When observables are incompatible, complex probabilities ($P(A \wedge B)$, $P(A \wedge \neg B)$, etc.) are measured in a characteristic ordering (indicated by the column headings). Differences in measurement ordering for simple and complex probabilities mean that the quantum probability model gives different predicted values for these identities when observables are compatible, when they are incompatible and measured in the order **AB**, or when they are incompatible and measured in the order **BA**. Observed average values for these identities are from Costello and Watts (2016a), Experiment 1 group 1 (5 pairs of A,B weather events: column C & W 1a), Costello and Watts (2016a), Experiment 1 group 2 (5 different pairs of weather events: column C & W 1b), Costello and Watts (2016a), Experiment 2 (6 pairs of future economic or political events: column C & W 2) and (Fisher & Wolfe, 2014) Experiments 1 and 2 (34 personality-description scenarios: column F & W). Note that Fisher and Wolfe (2014) did not gather estimates for identities 2,3,5 or 6.

Identity		Noise model	Quantum model			Observed values			
			Compatible	Incompatible		C & W 1a	C & W 1b	C & W 2	F & W
				Order **AB**	Order **BA**				
1	$P(A) + P(B) - P(A \wedge B) - P(A \vee B)$	0	0	δ_B	δ_A	0.00 (0.31)	0.03 (0.26)	−0.07 (0.28)	−0.03 (0.2)
2	$P(A \wedge B) + P(A \wedge \neg B) - P(A)$	d	0	0	$-\delta_A$	0.25 (0.31)	0.28 (0.32)	0.38 (0.29)	−
3	$P(A \wedge B) + P(\neg A \wedge B) - P(B)$	d	0	$-\delta_B$	0	0.24 (0.29)	0.20 (0.28)	0.43 (0.30)	−
4	$P(B|A)P(A) - P(A|B)P(B)$	0	0	$\delta_{A \wedge B}$	$\delta_{A \wedge B}$	−0.01 (0.12)	0.00 (0.16)	−0.03 (0.18)	−0.01 (0.12)
5	$P(A \wedge B) - P(A|B)P(B)$	$d/2$	0	$\delta_{A \wedge B}$	0	0.14 (0.18)	0.10 (0.18)	0.20 (0.21)	−
6	$P(A \wedge B) - P(B|A)P(A)$	$d/2$	0	0	$-\delta_{A \wedge B}$	0.13 (0.20)	0.10 (0.20)	0.17 (0.20)	−

exclusive subsets: the set of events $A \wedge B$, and the set of events $A \wedge \neg B$.

Consider an experiment where we ask people to estimate various probabilities $P(A)$, $P(B)$, $P(A \wedge B)$, $P(A \vee B)$, $P(A \wedge \neg B)$, $P(B \wedge \neg A)$ for a range of different events A and B, and combine those estimates as in the various identities. In such experiments we find two interesting patterns. First, when we combine people's probability estimates for a given pair of events A,B as in the addition law identity, the average value obtained is equal to, and symmetrically distributed around, probability theory's required value of 0. Second, when we combine the same estimates for the same events A,B as in the two expansion identities, the average value is not equal to 0; instead, the average value is positive and is similar for both of these expansion identities. In other words, people's probability estimates reliably agree with probability theory for the addition law identity, while simultaneously deviating from probability theory for the two expansion identities. Table 1, for example, shows average values for these identities from two experiments in Costello and Watts (2016a) and from a study by Fisher and Wolfe (2014): in this table we see that the average value for the addition law identity (identity 1 in the table) is very close to 0, while the average values for the two expansion identities (identities 2 and 3 in the table) are positive and have similar values.

This agreement with the addition law doesn't just hold when averaging across events: it also holds separately for each individual pair of events A and B, and holds when A and B are familiar everyday events (e.g. different types of weather), when A and B refer to the occurrence of different types of disease in certain populations, when A and B refer to future political or economic outcomes, or when A and B refer to various personality-description scenarios (Costello & Watts, 2014; Costello & Mathison, 2014; Costello & Watts,

Table 2

Means values for identities 1 and 4 for the pairs A and B used in Experiment 1 Group 1 (pairs 1–5), Experiment 1 Group 2 (pairs 6–10) and Experiment 2 (pairs 11–16) in Costello and Watts (2016a). Note that, for each individual pair, the mean value for these identities is close to 0 (the normatively correct value, and the value predicted by the noise model).

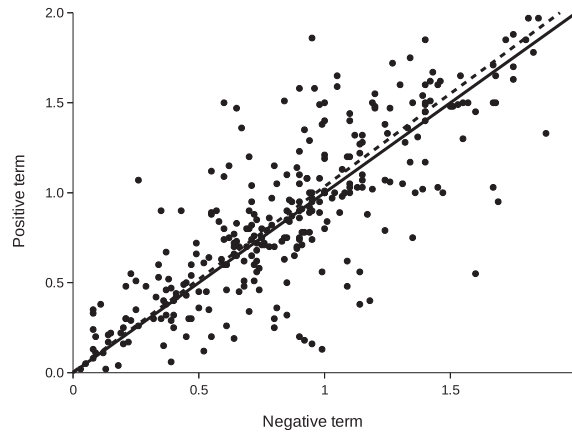| Pair			Identity	
	A	B	1 (addition law)	4 (Bayes rule)
1	Cold	Rainy	−0.07	0.05
2	Cloudy	Icy	0.08	−0.02
3	Cold	Thundery	−0.07	−0.02
4	Cloudy	Warm	0.03	−0.08
5	Sunny	Snowy	0.03	0.00
6	Cloudy	Rainy	−0.02	0.0
7	Cold	Icy	−0.07	−0.05
8	Cloudy	Thundery	0.04	−0.0
9	Sunny	Warm	−0.04	0.05
10	Icy	Snowy	−0.05	0.01
11	Britain has left the European Union	Greece has left the European Union	−0.04	−0.04
12	The number of cars on Irish roads has increased	The price of petrol has increased significantly	−0.05	−0.12
13	Climate change has a large impact on Ireland's weather	World greenhouse gas emissions have been reduced	−0.07	−0.06
14	The US is at war in the Middle East	There has been another major terrorist attack in the US	−0.03	0.01
15	Europe has grown noticably poorer	Unemployment in Europe is above 20%	−0.19	−0.07
16	Hurricanes and typhoons have become more frequent	The average world temperature has increased	−0.03	0.08
Overall			−0.03 (0.06)	−0.02 (0.06)

Fig. 2. Scatterplot of individual values for the positive and negative terms of the addition law identity, for all individual pairs A,B and all participants in Experiments 1 and 2 of Costello and Watts (2016a). Each point represents on the x-axis the sum of a single participant's estimates for $P(A \wedge B)$ and $P(A \vee B)$ for a single pair for events A,B (the negative terms in that identity), and on the y-axis the sum of the same participant's estimates for $P(A)$ and for $P(B)$ for the same events (the positive terms in that identity). There are 498 individual x,y pairs here in total: the correlation between x and y values (between positive and negative terms in the identity) was $r = 0.82$, $p < 0.00000001$. The dashed line represents the line of equality, and indicates the relationship required by normative probability theory. The line of best fit to the data (solid) was obtained via Deming regression, which accounts for error in observations on both the x- and y-axes.

2017, 2016a, 2016b; Fisher & Wolfe, 2014). To illustrate this, Table 2 shows average values for the addition law identity for each of the 16 different event pairs used in Costello and Watts (2016a): average values are close to 0 for almost all event pairs.

These results show that people's probability estimates, when put together in the form of the addition law identity, agree with the requirements of normative probability theory both when averaged across event pairs (Table 1) and when averaged with in a single event pair A and B (Table 2). This agreement holds despite reliable deviation from the requirements of normative probability theory for the two 'expansion' identities. These results do not, however, show that this agreement with the addition law holds in individual participant's responses. To demonstrate this agreement in individual responses, we constructed a scatterplot of individual values for the positive and negative terms of the addition law identity, for all individual pairs A,B and all participants in Experiments 1 and 2 of Costello and Watts (2016a). Each point in this scatterplot represents on the x-axis the sum of a single participant's estimates for $P(A \wedge B)$ and $P(A \vee B)$ for a single pair for events A,B (the negative terms in that identity), and on the y-axis the sum of the same participant's estimates for $P(A)$ and for $P(B)$ for the same events (the positive terms in that identity). If agreement with the addition law held reliably in participant's individual responses, we would expect points in this scatterplot to be distributed around the 'line of equality' (the line with intercept 0 and slope 1). Fig. 2 shows this scatterplot. As this figure shows, the relationship follows and is distributed symmetrically around the line of equality, indicating that the addition law identity does hold in individual probability estimates (with symmetrical deviations due to random error in individual responses). A JZS Bayes Factor analysis based on a paired t-test of x,y values in this scatterplot gave strong evidence in favour of the null hypothesis that x and y values were equal (Scaled JZS Bayes Factor = 19.84), supporting the conclusion that the addition law identity holds in individual participant probability estimates.

2.3. The Bayes rule identity

The addition law identity applies to direct or marginal probabilities. Similar results hold for identities that involve conditional probabilities. One identity involving conditionals is the additive form of Bayes rule, which requires that

$$P(B|A)P(A) - P(A|B)P(B) = 0 \tag{5}$$

must hold for all events A and B. Two parallel 'Bayes expansion' identities require that

$$P(A \wedge B) - P(A|B)P(B) = 0 \tag{6}$$

and

$$P(A \wedge B) - P(B|A)P(A) = 0 \tag{7}$$

must hold for all events A and B. These identities all follow from the conditional probability definition

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} \quad provided \ P(B) > 0$$

Again, consider an experiment where we ask people to estimate various probabilities $P(A)$, $P(B)$, $P(A \wedge B)$, $P(A|B)$, $P(B|A)$ and so on for a range of different events A and B, and combine those estimates as in the various identities. In such experiments we find patterns that follow those for the addition law identity. First, when we combine people's probability estimates as in the Bayes rule
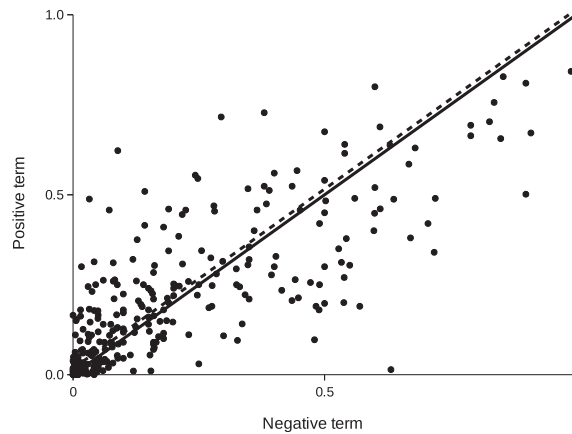
Fig. 3. Scatterplots of individual values for the positive and negative terms of the Bayes rule identity, from all participants in Experiments 1 and 2 of Costello and Watts (2016a). Each point represents on the *x*-axis the product of a single participant's estimates for $P(A|B)$ and $P(B)$ for a single pair for events A,B (the negative term in that identity), and on the *y*-axis the product of the same participant's estimates for $P(B|A)$ and $P(A)$ for the same events (the positive term). There are 498 individual *x,y* pairs here in total: the correlation between *x* and *y* values was $r = 0.81$, $p < 0.00000001$. The dashed line represents the line of equality, and indicates the relationship required by normative probability theory. The lines of best fit (solid) was obtained via Deming regression, which accounts for error in observations on both the *x*- and *y*-axes.

identity, the average value obtained is equal to, and symmetrically distributed around, probability theory's required value of 0. Second, when we combine the same estimates for the same events *A,B* as in the two Bayes expansion identities, the average value is not equal to 0; instead, the average value is positive (typically around 0.12, half the value seen for expansion identities in Eqs. (3) and (4)) and is similar for both of these expansion identities (see Table 1). Finally, as with the addition law identity, agreement with the Bayes rule identity holds both in averages across event pairs (see Table 1), in averages within individual event pairs (see Table 2). The identity also holds in individual responses, as demonstrated by the scatterplot in Fig. 3; as before, a JZS Bayes Factor analysis based on a paired *t*-test of *x,y* values in this scatterplot gave strong evidence in favour of the null hypothesis (Scaled JZS Bayes Factor = 19.85), supporting the conclusion that the identity holds in participant's probability estimates. Taken together, these results demonstrate that people's probability estimates reliably agree with probability theory for the Bayes rule identity, while simultaneously deviating from probability theory for the two Bayes expansion identities.

All of these patterns, of agreement with the addition law identity and the Bayes rule identity and simultaneous violation of the expansion identities (with approximately the same positive value for identities in Eqs. (3) and (4) and approximately half that value for identities in Eqs. (6) and (7)), are predicted to hold in the probability theory plus noise model. The model predicts similar patterns of agreement and bias for a number of similar identities (predictions which are confirmed in experimental results; see e.g. Costello & Watts, 2014, 2016a). Confirmation of these predictions has been taken as evidence that the probability theory plus noise model 'may provide a fully general account of the mechanisms by which people estimate probabilities' (Costello & Watts, 2016a).

3. The quantum probability model

The quantum probability model (Busemeyer, Pothos, Franco, & Trueblood, 2011; Busemeyer & Bruza, 2012; Wang & Busemeyer, 2013; Wang et al., 2014) assumes that people's probabilistic reasoning follows the mathematical rules used to calculate event probability in quantum theory. A fundamental aspect of quantum theory is that the probability of two quantum events can depend on the order in which those events are measured. This order dependence allows the quantum probability model to address various order effects seen in people's sequential inference and judgement.

Probability has a geometric interpretation in quantum theory, based on the projection of state vectors. We avoid this geometric interpretation here and instead focus on presenting various results on how quantum probability agrees with, and deviates from, standard probability theory. In quantum probability, an observable defines the set of all possible distinct outcomes for a given measurement: the set of possible answers to the question represented by that measurement. The primary theoretical distinction between quantum and standard probability lies in the idea of 'compatible' or 'incompatible' observables. Two observables are compatible if both observables can be measured simultaneously. If two observables are compatible, quantum probability theory reduces exactly to standard probability theory in all cases. This means that if two observables are compatible then all the probability theory identities described above have a value of 0, and there are no order effects in judgement.

Incompatible observables, by contrast, cannot be measured simultaneously, and measurement outcomes depend on the order of measurement. If all probabilities are measured with the same ordering then again quantum probability theory reduces exactly to standard probability theory (if all probabilities are of the form $P_{AB}()$, for example, then all relationships between those probabilities match the requirements of standard probability theory and all probability theory identities hold). If probabilities are measured with different orderings, however, then quantum probability deviates from standard probability, producing biases in judgement and order

effects in sequential question answering such as $P_{\mathbf{AB}}(A) \neq P_{\mathbf{BA}}(A)$, $P_{\mathbf{BA}}(B) \neq P_{\mathbf{AB}}(B)$ and $P_{\mathbf{AB}}(A \wedge B) \neq P_{\mathbf{BA}}(A \wedge B)$. In our presentation of the quantum probability model we represent these order effects as

$$\delta_A = P_{\mathbf{AB}}(A) - P_{\mathbf{BA}}(A)$$
$$\delta_B = P_{\mathbf{BA}}(B) - P_{\mathbf{AB}}(B)$$
$$\delta_{A \wedge B} = P_{\mathbf{AB}}(A \wedge B) - P_{\mathbf{BA}}(A \wedge B)$$

Each of these represents a 'quantum interference' term for A,B or $A \wedge B$. These interference terms arise from the geometric relationship between the current state vector, the vector representing the observable **A** and that representing the observable **B**. Note that these quantum interference expressions are not error terms: for given observables **B** and **A** (and a given state, in quantum theory; or a given participant, in the quantum probability model) these terms have a fixed values that represented the difference between probabilities when measured in different orders. The quantum interference term δ_A, for example, has a fixed value that specifies the relationship between $P_{\mathbf{BA}}(A)$ (the probability of A when estimated immediately after presentation of B) and $P_{\mathbf{AB}}(A) = P_*(A)$ (the probability of A when estimated without prior presentation of B). These quantum interference terms can take on different values for different observables (and different participants): in some cases positive, in some negative, and in some cases 0.

It is important at this point to make a distinction between quantum probability applied as a model of choice (a model of how people choose to answer 'yes' or 'no' to consecutive questions in, for example, QQ identity question tasks) and quantum probability applied as a model of probability estimation (a model of how people produce an estimate of the probability of a given event, as in the addition law and the Bayes rule tasks). In applying the quantum model to a choice situation involving consecutive questions **A** and **B** (as in QQ identity question tasks), theorists take quantum probabilities such as $P_{\mathbf{AB}}(A)$ to represent the probability of a participant answering 'yes' to question **A** (see e.g. Wang & Busemeyer, 2013; Wang et al., 2014). Here the quantum probability represents the chance of a given response, rather than a participant's probabilistic judgement about the event in question: so that $P_{\mathbf{AB}}(A)$ represents the probability of a participant answering 'yes' to a question about Clinton being trustworthy when that question is asked first, but $P_{\mathbf{AB}}(A)$ does not represent a specific participant's individual estimate of the probability that Clinton is trustworthy. When taking this view, quantum probabilities such as $P_{\mathbf{AB}}(A)$ are in some way 'ensemble' properties: they apply to the set of participants taking part in a given question–answering task, rather than to each participant individually.

In applying the quantum model to situations involving probability judgement (that is, situations where participants are asked to estimate the probability of some event A), by contrast, theorists assume that the quantum probability $P_{\mathbf{AB}}(A)$ *does* represent an average participant's estimate or judgement of the probability of A. We can see this aspect of the quantum model most clearly in the model's application to the well-known conjunction fallacy, which arises when individual participants reliably judge the probability of a conjunctive event $A \wedge B$ to be higher than the probability of a constituent event A (Tversky & Kahneman, 1983). One argument in favour of the QP approach has been its ability to explain this fallacy in terms of differential quantum interference effects for the single probability $P_{\mathbf{AB}}(B)$ and the conjunctive probability $P_{\mathbf{AB}}(A \wedge B)$: a result that can only hold if quantum probabilities such as $P_{\mathbf{AB}}(B)$ and $P_{\mathbf{AB}}(A \wedge B)$ are assumed to represent people's probability estimates, rather than the chance of a given response in a choice situation (see e.g. Busemeyer et al., 2011). We see a similar approach in research investigating order effects in probability judgement, where participants are given, for example, a sequence of facts about a criminal case and asked to estimate, after each new piece of information, the probability of the defendant being guilty. The quantum probability model gives an accurate computational fit to differences in people's probability estimates depending on the order of presentation of information in such studies: a result that can only hold if, as before, quantum probabilities are assumed to represent people's probability estimates (see, e.g. Trueblood & Busemeyer, 2011). Finally, Yearsley and Trueblood's recent work makes this link explicit, by using the quantum model to predict agreement with a version of the QQ identity in subjective probability estimates (as discussed above).

This link between probability in choice tasks and probability in estimation tasks is, of course, psychologically plausible (it is reasonable to think that the chance of someone responding 'yes' to a question about a given statement in a choice task is proportional to their judgement of the likelihood of that statement being true). We stress this link explicitly here because it is necessary for our application of the quantum model simultaneously to the QQ, addition law, and Bayes rule identities, below, since the first identity arises in choice tasks, while the second and third arise in probability estimation tasks. We return to this topic in the General Discussion below.

3.1. The addition law

The addition law identity applies in cases where questions are not presented in a particular sequential order (**AB** or **BA**) but are order independent. In this situation there are no order effects for simple probabilities (the probability of A is $P_*(A) = P_{\mathbf{AB}}(A)$ and that of B is $P_*(B) = P_{\mathbf{BA}}(B)$). Since in quantum probability incompatible observables must be measured in a particular order, however, order effects for incompatible observables still apply when people are asked to estimate conjunctive or disjunctive probabilities such as $P(A \wedge B)$, $P(A \wedge \neg B)$ or $P(A \vee B)$. For such conjunctions or disjunctions the quantum probability model assumes a particular characteristic ordering for observables that depends on the causal link between those observables, or possibly on the values of the simple probabilities $P(A)$ and $P(B)$. Complex probabilities such as $P(A \wedge B)$ are estimated using this characteristic ordering. This means that the relationship between a simple probability $P(A)$ and the conjunctive probabilities $P(A \wedge B)$ and $P(A \wedge \neg B)$ will depend on this characteristic ordering. When the characteristic ordering of observables for conjunctions is **AB**, the relationship between estimates for $P(A)$ and estimates for $P(A \wedge B)$ and $P(A \wedge \neg B)$ will be

$$P_*(A) = P_{\mathbf{AB}}(A) = P_{\mathbf{AB}}(A \wedge B) + P_{\mathbf{AB}}(A \wedge \neg B) \tag{8}$$

as in standard probability theory (since the ordering of observables is the same for all three probabilities in this expression, quantum probability reduces to standard probability in this case). When the characteristic order of observables for conjunctions is **BA**, however, the relationship between estimates for $P(A)$ and estimates for $P(A \wedge B)$ and $P(A \wedge \neg B)$ will be

$$P_*(A) = P_{\mathbf{AB}}(A) = P_{\mathbf{BA}}(A) + \delta_A = P_{\mathbf{BA}}(A \wedge B) + P_{\mathbf{BA}}(A \wedge \neg B) + \delta_A \tag{9}$$

with the second equality arising from the definition of δ_A, and the third equality arising from the fact that $P_{\mathbf{BA}}(A)$, $P_{\mathbf{BA}}(A \wedge B)$ and $P_{\mathbf{BA}}(A \wedge \neg B)$ are all measured in the same ordering (and so the requirements of standard probability theory hold and so $P_{\mathbf{BA}}(A) = P_{\mathbf{BA}}(A \wedge B) + P_{\mathbf{BA}}(A \wedge \neg B)$). Parallel results, of course, arise for the probability of B: when complex probabilities are measured in the ordering **BA** we get

$$P_*(B) = P_{\mathbf{BA}}(B) = P_{\mathbf{BA}}(A \wedge B) + P_{\mathbf{BA}}(\neg A \wedge B) \tag{10}$$

as in standard probability theory, but when complex probabilities are measured in the ordering **AB** we get

$$P_*(B) = P_{\mathbf{BA}}(B) = P_{\mathbf{AB}}(B) + \delta_B = P_{\mathbf{AB}}(A \wedge B) + P_{\mathbf{AB}}(\neg A \wedge B) + \delta_B \tag{11}$$

We can use these expressions for $P_*(A)$ and $P_*(B)$ to derive the quantum probability model's predictions for values of the addition law identity (Eq. (2)) and for the two expansion identities (Eqs. (3) and (4)) in three separate situations: where observables are compatible, where the characteristic ordering of observables is **AB**, and where the characteristic ordering is **BA**. The first three lines of Table 1 shows these predictions. Recall that experimental results show that the addition law identity has a value of approximately 0 in people's probability judgements, while the two expansion identities have approximately the same positive value. From Table 1 we see that, if observables are compatible, all three identities have a predicted value of 0 (contrary to experimental results). If observables are measured in the order **AB**, one expansion identity has a predicted value of 0 and the addition law and the other expansion identity have the same values, deviating from zero by δ_A (contrary to experimental results). If observables are measured in the order **BA**, the other expansion identity has a predicted value of 0 and the addition law and the first expansion identity have the same values, deviating from zero by δ_B (again, contrary to experimental results). The quantum probability model's predictions are inconsistent with the experimental results in all three situations: the quantum model cannot account for these results.

3.2. Bayes rule identity

In quantum probability theory a conditional probability $P(A|B)$ is necessarily measured in the order **BA** (with the given event B occurring first and the conditional event A occurring after). This means that the relationships

$$P_{\mathbf{BA}}(A \wedge B) = P(A|B)P_{\mathbf{BA}}(B) = P(A|B)P_*(B) \tag{12}$$

$$P_{\mathbf{AB}}(A \wedge B) = P(B|A)P_{\mathbf{AB}}(A) = P(B|A)P_*(A) \tag{13}$$

necessarily hold in quantum probability (since the probabilities in these expressions are all measured in the same order, and so follow the requirements of probability theory). If observables **A** and **B** are compatible, there are no order effects and so we have

$$P(B|A)P_*(A) - P(A|B)P_*(B) = P_{\mathbf{BA}}(A \wedge B) - P_{\mathbf{AB}}(A \wedge B) = 0$$

as in standard probability theory. If observables are incompatible, however, we get

$$P(B|A)P_*(A) - P(A|B)P_*(B) = P_{\mathbf{BA}}(A \wedge B) - P_{\mathbf{AB}}(A \wedge B) = \delta_{A \wedge B}$$

The quantum probability model thus predicts a value of 0 for the Bayes rule identity only when observables are compatible (that is, when there are no order effects), while for incompatible observables, the quantum model predicts a value for the Bayes rule identity equal to $\delta_{A \wedge B}$.

We can derive similar predictions for the 'Bayes expansion' identities (Eqs. (6) and (7)) in three separate situations: where observables are compatible, where the characteristic ordering of observables is **AB**, and where the characteristic ordering is **BA**. The last three lines of Table 1 shows the predictions for these three identities. Recall that experimental results show that the Bayes rule identity has a value of approximately 0 in people's probability judgements, while the two Bayes expansion identities have approximately the same positive value. From Table 1 we see that, if observables are compatible, all three identities have a predicted value of 0 (contrary to experimental results). If observables are measured in the order **AB**, one expansion identity has a predicted value of 0 and the Bayes rule and the other expansion identity have the same values, deviating from zero by $\delta_{A \wedge B}$ (contrary to experimental results). If observables are measured in the order **BA**, the second Bayes expansion identity has a predicted value of 0 and the Bayes rule and the other expansion identity have values that both deviate from zero by $\delta_{A \wedge B}$ but with opposite signs (again, contrary to experimental results). The quantum probability model's predictions are inconsistent with the experimental results in all three situations: again, the quantum model cannot account for these results.

3.3. The 'QQ' identity

Unlike the addition law and Bayes rule identities, the QQ identity applies specifically to cases where questions are presented in a given sequential order (**AB** or **BA**). The order effects for incompatible observables in the quantum probability model imply that if participants are asked to answer sequential questions about incompatible observables **A** and **B**, then $P_{\mathbf{AB}}(A \wedge B)$ (the probability of a 'yes' answer to both questions when the order is **AB**) will be different from the probability $P_{\mathbf{BA}}(A \wedge B)$ (the probability of a 'yes'

answer to both questions when the order is **BA**), with this difference being represented by the interference term $\delta_{A \wedge B}$. A necessary mathematical consequence of quantum probability is that exactly the same order effects apply to conjunctive probabilities $P_{\mathbf{BA}}(\neg A \wedge \neg B)$ and $P_{\mathbf{AB}}(\neg A \wedge \neg B)$, and so we have

$$P_{\mathbf{AB}}(A \wedge B) - P_{\mathbf{BA}}(A \wedge B) = \delta_{A \wedge B} = P_{\mathbf{BA}}(\neg A \wedge \neg B) - P_{\mathbf{AB}}(\neg A \wedge \neg B)$$

and therefore the QQ identity holds for events A and B in the quantum probability model (see Wang & Busemeyer, 2013; Wang et al., 2014, for proofs). By symmetry we have parallel agreement between order effects

$$P_{\mathbf{AB}}(A \wedge \neg B) - P_{\mathbf{BA}}(A \wedge \neg B) = P_{\mathbf{BA}}(\neg A \wedge B) - P_{\mathbf{AB}}(\neg A \wedge B)$$

Wang et al. (2014) estimate the size of the overall order effect in each of their 72 different polls or experimental studies via the overall order effect measure Z, where

$$Z = max \begin{cases} |P_{\mathbf{BA}}(A \wedge B) - P_{\mathbf{AB}}(A \wedge B)| + |P_{\mathbf{BA}}(\neg A \wedge \neg B) - P_{\mathbf{AB}}(\neg A \wedge \neg B)| \\ |P_{\mathbf{BA}}(A \wedge \neg B) - P_{\mathbf{AB}}(A \wedge \neg B)| + |P_{\mathbf{BA}}(\neg A \wedge B) - P_{\mathbf{AB}}(\neg A \wedge B)| \end{cases} \tag{14}$$

The overall order effect for a given survey, as measured by this expression Z, is equal to the summed absolute values of the order effects for $A \wedge B$ and $\neg A \wedge \neg B$, or for $A \wedge \neg B$ and $\neg A \wedge B$, whichever is greater. The greater the value of this measure for a given survey, the larger the overall order effect in that survey. Wang et al. (2014) find statistically significant order effects in most of these polls or studies, but reliable agreement with the QQ identity. The fact that this QQ identity appears to hold simultaneously with such order effects has been taken as clear evidence that 'human judgements follow quantum rules' (Wang et al., 2014). In general, however, it seems that the quantum probability model accounts for satisfaction of the QQ identity alongside the occurrence of reliable order effects in sequential judgement, but is unable to account for addition law and Bayes law results.

4. The probability theory plus noise model

The probability theory plus noise model assumes that people estimate probabilities via a mechanism that is fundamentally rational (following standard frequentist probability theory), but is perturbed in various ways by the systematic effects or biases caused by purely random noise or error. This approach follows a line of research leading back at least to Thurstone (1927) and continued by various more recent researchers (see, e.g. Dougherty, Gettys, & Ogden, 1999; Erev, Wallsten, & Budescu, 1994; Hilbert, 2012). This model explains a wide range of results on bias in people's direct and conditional probability judgements across a range of event types, and identifies various probabilistic expressions in which this bias is 'cancelled out' and for which people's probability judgements agree with the requirements of standard probability theory (see Costello & Watts, 2014, 2017, 2016a, 2016b; Costello & Mathison, 2014).

In standard frequentist probability theory the probability of some event A is estimated by drawing a random sample of events, counting the number of those events that are instances of A, and dividing by the sample size. The expected value of these estimates is $P(A)$, the probability of A; individual estimates will vary with a binomial proportion distribution around this expected value. Our model assumes that people estimate the probability of some event A in exactly the same way: by randomly sampling items from memory, counting the number that are instances of A, and dividing by the sample size. If this process was error-free, people's estimates would be expected to have an average value of $P(A)$ (and to vary randomly around that average, due to sampling error). Human memory is subject to various forms of random error, however. To reflect this we assume events have some chance $d < 0.5$ of randomly being read incorrectly: there is a chance d that a $\neg A$ (*not A*) event will be incorrectly counted as A, and the same chance d that an A event will be incorrectly counted as $\neg A$. We take $P_*(A)$ to represent $P(read\ as\ A)$: the probability that a single randomly sampled item from this population will be read as an instance of A (subject to this random error in counting). Since a randomly sampled event will be counted as A if the event truly is A and is counted correctly (this occurs with a probability $(1-d)P(A)$, since $P(A)$ events are truly A and events have a $1-d$ chance of being counted correctly), or if the event is truly $\neg A$ and is counted incorrectly as A (this occurs with a probability $(1-P(A))d$, since $1-P(A)$ events are truly $\neg A$, and events have a d chance of being counted incorrectly), the population probability of a single randomly sampled item being read as A is

$$P(read\ as\ A) = P_*(A) = (1-d)P(A) + (1-P(A))d = (1-2d)P(A) + d \tag{15}$$

We now consider the process of probability estimation. We take $p_*(A)$ to represent an individual estimate of the probability of A, produced by randomly sampling some set of events from memory and counting the proportion that are A (subject to random error in reading an item as A). Since $p_*(A)$ is the probability of an item being read as A, and since these samples are drawn randomly, these estimates $p_*(A)$ will vary randomly following the binomial proportion distribution

$$\frac{Bin(N, P_*(A))}{N}$$

where N is the size of the sample drawn. Using $\langle X \rangle$ to represent the expected value of some randomly varying variable X (the value we would get if we averaged an infinite number of samples of that variable), a property of the binomial proportion distribution is that

$$\left\langle \frac{Bin(N, P_*(A))}{N} \right\rangle = P_*(A)$$

for any sample size N. Given this, we take $P_*(A)$ to represent the expected value of estimates $p_*(A)$ independent of sample size: the

value we would get if we averaged an infinite number of individual estimates $p_*(A)$, each based on a sample drawn randomly from a population with probability $P(A)$ noise rate d, and with sample size varying across samples. Let p_i represent the probability of a sample being drawn with a particular size $N = i$, and we have

$$\langle p_*(A) \rangle = \sum_{i=1}^{\infty} p_i \left\langle \frac{Bin(i, P_*(A))}{i} \right\rangle = \sum_{i=1}^{\infty} p_i P_*(A) = P_*(A) \sum_{i=1}^{\infty} p_i$$

Since the sum of probabilities p_i across all sample sizes necessarily equals 1, we thus have

$$\langle p_*(A) \rangle = P_*(A) = (1-2d)P(A) + d \tag{16}$$

This equation gives the expected value or predicted average for people's estimates for the probability of some event A: individual estimates will vary randomly around this expected value in a binomial proportion distribution. Note that this predicted average embodies a regression towards the center, due to random noise: estimates are systematically biased away from the 'true' probability $P(A)$, such that on average estimates will tend to be greater than $P(A)$ when $P(A) < 0.5$, and will tend to be less than $P(A)$ when $P(A) > 0.5$, and will tend to equal $P(A)$ when $P(A) = 0.5$. This regression due to noise accounts for a number of observed patterns of systematic bias in people's probability estimates, such as underconfidence, subadditivity, the conjunction fallacy, and the disjunction fallacy (see Costello & Watts, 2014, 2017, 2016b).[2]

This model can also be extended to the estimation of conditional probabilities $P(A|B)$ (the probability of A given that B has occurred). Reasoning just as above, the model predicts an expected value for the conditional probability $P(A|B)$ of

$$P_*(A|B) = \frac{(1-2d)^2 P(A \wedge B) + d(1-2d)[P(A) + P(B)] + d^2}{(1-2d)P(B) + d} \tag{17}$$

and accounts for various observed patterns of bias in people's conditional probability judgement (see Costello & Watts, 2016a for the derivation of this expression).

4.1. Addition law and Bayes rule identities

As well as accounting for various patterns of systematic bias, this model also makes predictions about the values of various probability theory identities such as the addition law and the Bayes rule identity. If we substitute the expected values from Eq. (15) into the addition law identity, for example, we get an expected value of

$$\begin{aligned} &P_*(A) + P_*(B) - P_*(A \wedge B) - P_*(A \vee B) \\ &= (1-2d)P(A) + d + (1-2d)P(B) + d \\ &\quad - (1-2d)P(A \wedge B) - d - (1-2d)P(A \vee B) - d \\ &= 0 \end{aligned}$$

and so this model predicts that this expression should have a value of 0 on average in people's probability judgements just as required by standard probability theory, and just as seen in experimental results (Costello & Watts, 2016a, 2016b, 2014, 2017; Costello & Mathison, 2014; Fisher & Wolfe, 2014). Similarly, if we substitute the expected values from Eqs. (15) and (17) into the Bayes rule identity, we get an expected value of

$$\begin{aligned} &P_*(B|A)P_*(A) - P_*(A|B)P_*(B) \\ &= (1-2d)^2 P(A \wedge B) + d(1-2d)[P(A) + P(B)] + d^2 \\ &\quad - (1-2d)^2 P(A \wedge B) + d(1-2d)[P(B) + P(A)] - d^2 \\ &= 0 \end{aligned}$$

and again the model predicts a value of 0, just as required by standard probability theory and just as seen in experimental results (Costello & Watts, 2016a; Fisher & Wolfe, 2014).

Note that for both the addition law and the Bayes rule identity, the probability theory plus noise model predicts an average value of 0, as required by normative probability theory, with individual values varying randomly and symmetrically around that value, due to random noise. This is just the pattern we see with individual values for these identities (see Figs. 2 and 3). The model does not predict that every individual value for these identities will equal 0 (this would be a very unreasonable prediction for a model which gives a fundamental role to random noise). Agreement with probability theory for the addition law and the Bayes rule identity arises in this model despite significant regressive bias due to random noise in individual probability estimates making up these expressions. This is because in the addition law and Bayes rule expressions the various biases due to random noise in those individual probability estimates all cancel out, leaving no overall bias in the identity as a whole. For other probability theory identities, however, this model predicts no cancellation of regressive effects. For example, substituting the expected value expressions from our model into the first 'expansion' identity (Eq. (3)), we get

---

[2] In accounting for effects such as the conjunction fallacy, the noise model makes use of the standard statistical assumption of 'propagation of error', which says that noise rates for conjunctions and disjunctions are slightly higher than rates for single events (rates of d for single events, rates of $d + \Delta d$ for conjunctions and disjunctions; Δd assumed to be small). For simplicity of presentation we don't use this Δd term here: since Δd is assumed to be small, it has very little influence on the model's predictions about values of the various identities. We describe the effect of this Δd term on the addition law and Bayes rule identities in detail in Costello and Watts (2016a).

$$P_*(A \wedge B) + P_*(A \wedge \neg B) - P_*(A) = (1-2d)P(A \wedge B) + d] + (1-2d)P(A \wedge \neg B) + d - (1-2d)P(A) - d$$
$$= d$$

and, since the error term d is necessarily positive, our model predicts that values for this identity, computed from people's estimates, will have a reliably positive value. For the second 'expansion' identity (Eq. (4)) we similarly get

$$P_*(A \wedge B) + P_*(\neg A \wedge B) - P_*(B) = (1-2d)P(A \wedge B) + d] + (1-2d)P(\neg A \wedge B) + d - (1-2d)P(B) - d$$
$$= d$$

and the model predicts the same positive value for both identities, again just as observed in experimental results (Costello & Watts, 2014, 2017; Costello & Mathison, 2014 and see Table 1).

For the two 'Bayes expansion' identities (Eqs. (6) and (7)) we get

$$P_*(A \wedge B) - P_*(A|B)P_*(B) = (1-2d)P(A \wedge B) + d - (1-2d)^2 P(A \wedge B)$$
$$- d(1-2d)[P(A) + P(B)] - d^2$$
$$= d(1-d) - d(1-2d)[P(A) + P(B) - 2P(A \wedge B)]$$

and

$$P_*(A \wedge B) - P_*(B|A)P_*(A) = (1-2d)P(A \wedge B) + d - (1-2d)^2 P(A \wedge B)$$
$$- d(1-2d)[P(A) + P(B)] - d^2$$
$$= d(1-d) - d(1-2d)[P(A) + P(B) - 2P(A \wedge B)]$$

Since probability theory requires that $0 \leqslant P(A) + P(B) - 2P(A \wedge B) \leqslant 1$ for all A and B, and since $d < 0.5$ by assumption, we see that

$$d^2 \leqslant d(1-d) - d(1-2d)[P(A) + P(B) - 2P(A \wedge B)] \leqslant d(1-d)$$

and values for both these identities will be distributed between d^2 and $d(1-d)$ in a way that depends on $P(A) + P(B) - 2P(A \wedge B)$. The average value for $P(A) + P(B) - 2P(A \wedge B)$ (across uniformly distributed probabilities that are constrained to be consistent with probability theory) is 0.5, and so the average value for this expression is equal to $d/2$, the centerpoint of this range. The model thus predicts the same average positive value for both these identities; a value that is half that for the first two expansion identities. Again, this is just as seen in experimental results (Costello & Watts, 2016a and Table 1).

While the probability theory plus noise model can account for the observed addition law, Bayes rule identity, and expansion identity results, the model gives no role to order in probability estimation and so is unable to account for the QQ identity results. In the next section we describe how this model can be extended to account for sequential order effects in probability estimation, and show that this extension accounts both for systematic bias due to order in probability estimates, and for simultaneous agreement with the QQ identity.

4.2. The QQ identity and order effects

The probability theory plus noise model, as presented above, assumes that in sequential probability judgements where one judgement $P(B)$ is immediately followed by another judgement $P(A)$, the value given for $P(A)$ is not influenced by the prior value given for $P(B)$. Why does the noise model assume no influence between sequential probability judgements? The basic cause is the assumption that people estimate some probability $P(A)$ by drawing a sample of items *at random* from memory, and counting the proportion that are A. To allow sequential effects in the noise model, we can relax this assumption, and say that the chance of a given item being sampled from memory is influenced by the degree to which that item is already active: more active items in memory are primed, and are more likely to be 'randomly' sampled than the less active items. Assume that probabilities are estimated in the order **B** then **A**. Since the estimation of probability $P(B)$ involved drawing a sample of items and counting the proportion that are B, those items that were counted as B are more active (are primed), and so are more likely to be included in the 'random' sample of items drawn when estimating $P(A)$. This idea, that priming affects the sample drawn when estimating a probability, is consistent with the idea that probability judgements are systematically influenced by various forms of random noise or error: priming, here, becomes one particular form of random error that causes judgements to vary from estimate to estimate.

Suppose that the chance of an already primed item being sampled is s. Also suppose that $P(B)$ has just been estimated in a previous sample: $P_*(B)$ then represents the proportion of items in that previous sample that were read as B. A sample is now drawn to estimate $P(A)$. Each item drawn to make up that new sample has a probability $sP_*(B)$ of coming from the primed set of items that were already read as B, and a probability $1 - sP_*(B)$ of being drawn randomly from the set of all items in memory. For the $sP_*(B)$ items in our sample that were previously read as B, the probability of one of those items being read as A is $P_*(A|B)$; this is the conditional probability of an item being read as A, given that it was read as B. For the remaining items that were just sampled randomly from memory, the probability of one of those items being read as A is simply $P_*(A)$. Given that we have just given an estimate for the probability $P(B)$, then, the expected value for an immediately following estimate for $P(A)$ will be

$$P_{\mathbf{BA}}(A) = sP_*(B)P_*(A|B) + (1 - sP_*(B))P_*(A) \tag{18}$$

and, substituting from Eqs. (15) and (17) and simplifying we get

$$P_{\mathbf{BA}}(A) = P_*(A) + s(1-2d)^2[P(A \wedge B) - P(A)P(B)] \tag{19}$$

as the probability of answering 'yes' to a question **A** when that question comes immediately after a question **B**.

From Eq. (19) we see that $P_{\mathbf{BA}}(A) \neq P_*(A)$ and so $P_{\mathbf{BA}}(A) \neq P_{\mathbf{AB}}(A)$ will hold in this model in general, with the probability of a 'yes' answer to question **A** when that question comes first being different from the probability of a 'yes' answer when question **A** immediately follows question **B**. This model thus produces order effects in question answering, just as seen in experimental data. The sequential order effect here is proportional to the difference $P(A \wedge B) - P(A)P(B)$. This difference will be positive if B is positively associated with A (if A is more likely to occur given that B has occurred), will be negative if A is negatively associated with B (if A is less likely to occur given B has occurred), and there will be no such contextual effect if A and B are independent (if the occurrence of B has no impact on the probability of occurrence of A).

Despite these order effects, the QQ identity also holds in this model. To see this, consider that, since $P_*(B)$ is the probability of answering 'yes' to a question **B** and $P_{\mathbf{BA}}(A)$ is the probability of answering 'yes' to a question **A** that immediately follows a question **B**, the probability of answering 'yes' to both questions when presented in the order **BA** is

$$P_{\mathbf{BA}}(A \wedge B) = P_*(B)P_{\mathbf{BA}}(A) = P_*(B)P_*(A) + P_*(B)s(1-2d)^2[P(A \wedge B) - P(B)P(A)]$$

and the probability of answering 'yes' to both questions in the order **AB** is

$$P_{\mathbf{AB}}(A \wedge B) = P_*(A)P_{\mathbf{AB}}(B) = P_*(A)P_*(B) + P_*(A)s(1-2d)^2[P(A \wedge B) - P(B)P(A)]$$

and so

$$P_{\mathbf{BA}}(A \wedge B) - P_{\mathbf{AB}}(A \wedge B) = s(1-2d)^2[P(A \wedge B) - P(B)P(A)][P_*(B) - P_*(A)] \qquad (20)$$

Using the same line of reasoning for the probability of answering 'no' to both questions, we get

$$P_{\mathbf{AB}}(\neg B \wedge \neg A) - P_{\mathbf{BA}}(\neg B \wedge \neg A) = s(1-2d)^2[P(\neg B \wedge \neg A) - P(\neg B)P(\neg A)][P_*(\neg A) - P_*(\neg B)]$$

Substituting from Eq. (15) and rearranging we have

$$P_*(\neg A) - P_*(\neg B) = (1-2d)[1-P(A)] + d - (1-2d)[1-P(B)] - d \ = P_*(B) - P_*(A)$$

and from standard probability theory we have

$$P(\neg A \wedge \neg B) - P(\neg A)P(\neg B) = P(A \wedge B) - P(B)P(A)$$

and so

$$P_{\mathbf{AB}}(\neg A \wedge \neg B) - P_{\mathbf{BA}}(\neg A \wedge \neg B) = s(1-2d)^2[P(B \wedge A) - P(B)P(A)][P_*(B) - P_*(A)] \qquad (21)$$

giving

$$P_{\mathbf{AB}}(B \wedge A) - P_{\mathbf{BA}}(B \wedge A) = -[P_{\mathbf{AB}}(\neg B \wedge \neg A) - P_{\mathbf{BA}}(\neg B \wedge \neg A)]$$

and this model satisfies the QQ identity. Note, however, that this derivation of the QQ identity is in terms of probability estimates: it assumes that people's sequential estimates of probability will be subject to priming and cancellation effects such that the QQ identity holds, on average, but does not say anything explicit about people's responses in choice tasks. To apply this model to the situation of sequential choice (sequential 'yes/no' answers, as in Wang et. al.'s QQ identity results), we link probability estimation and choice by assuming that, to a first approximation, the chance of a given participant answering 'yes' to a given question statement is simply proportional to that participant's judgement of the probability of that statement being true. Given this, the model predicts satisfaction of the QQ identity in both probability estimation and choice tasks.

How does this account of the effect of priming on probability judgement in sequentially ordered tasks connect to our earlier model of noisy probability estimation? This account represents a more detailed description of one particular aspect of noise affecting probability judgement. In our model we assume that noise in probability judgement has many different sources: random error in counting, random error in sampling, response error, and, indeed, priming effects: when items are presented in randomly varying order, priming effects are simply another source of random variation. In situations where probability estimation questions are presented in random order, we represent all these sources of random noise in terms of a single noise parameter d, as in the model presented earlier. When items are presented in a particular sequential order, however, these priming effects are no longer random but are controlled by the order of presentation; in such situations we take the parameter d to represent the various remaining random effects, and use our priming parameter s to represent the, now no longer random, effect of order on probability estimation. This account of priming thus does not, strictly speaking, represent a modification of the probability theory plus noise model; instead it simply represents a more detailed elaboration of one specific factor influencing noise in probability judgement.

Given this account of priming in sequential order effects, the noise model can now account for agreement with the addition law and Bayes rule identities in people's probability estimation, alongside simultaneous violation of the two expansion identities in those judgements, and can also account for agreement with the QQ identity in people's sequential question answering, alongside simultaneous order effects in people's question answers. This model thus gives a general account for patterns of agreement with, and deviation from, standard probability theory in probability estimation and in sequential question answering.

It is worth noting that we are not alone in presenting an alternative account of the QQ identity that does not depend on quantum probability effects: in a recent paper Kellen, Singmann, and Batchelder (2017) described a class of repeat-choice models that also predict the QQ identity. In the next section we test a specific prediction derived from our priming-based account of order effects in sequential question answering.

4.3. Predicting order effects in sequential question answering

Order effects arise when people's answers for a given question are influenced by the order of question presentation: when the probability of a 'yes' answer to question **A** when that question comes first is significantly different from the probability of a 'yes' answer when question **A** is preceded by question **B** ($P_{\mathbf{AB}}(A) \neq P_{\mathbf{BA}}(A)$). The probability theory plus noise model makes a specific prediction about the direction and magnitude of such order effects. Rearranging Eq. (18) we get

$$P_{\mathbf{BA}}(A) - P_{\mathbf{AB}}(A) = P_{\mathbf{BA}}(A) - P_*(A) = sP_*(B)P_*(A|B) - sP_*(B)P_*(A) = sP_*(B)[P_*(A|B) - P_*(A)] \tag{22}$$

(where, as before, s is the chance of an already primed item being sampled). Note that the probability term on the right-hand-side here measures, in classical probability theory, the degree of dependence between A and B. If we make the simplifying first-order assumption that the parameter s is constant across questions, the noise model thus predicts that the effect of question order on responses ($P_{\mathbf{BA}}(A) - P_*(A)$) will be reliably correlated, across questions, with degree of dependence between A and B, $P_*(B)[P_*(A|B) - P_*(A)]$. There will be no order effect, in the noise model, if A and B are independent (if $P_*(A|B) = P_*(A)$). There will be a positive order effect (a 'yes' answer to question **A** being more likely when presented after question **B** than otherwise) when A and B are positively dependent, and a negative order effect (a 'yes' answer to question **A** being less likely after question **B** than otherwise) when A and B are negatively dependent.

The quantum probability model, by contrast, doesn't predict any specific direction or magnitude for order effects. In the quantum model the order effect for A follows the expression

$$P_{\mathbf{BA}}(A) - P_{\mathbf{AB}}(A) = -\delta_A = 2P_{\mathbf{BA}}(A \wedge B) - 2\theta\sqrt{P_*(B)}\sqrt{P_*(A)}$$

for a parameter θ that depends on the relationship between the geometric vectors assumed to represent A,B, and the current state (see Wang & Busemeyer, 2013, p. 697). This parameter θ is *a priori* not known, is different for each pair of questions A and B, and can only be estimated in terms of observed values for the order effect for those specific questions (see Wang & Busemeyer, 2013, p. 698). This parameter cannot be used to predict the sign or magnitude of the order effect (since its value must be estimated in terms of that effect).

We can test the noise model's prediction about the order effect in question answering using the same dataset originally used by Wang et al. (2014) to show agreement with QQ identity.[3] This dataset gathers people's sequential question answering responses across a wide range of different topics in 72 different national representative surveys in the US (with responses from typically around 800 participants per topic), and in laboratory studies of the effects of order in question answering (Wang et al., 2014). In this dataset we measure the order effect for questions in each individual survey by calculating the difference between the probability of a 'yes' answer to survey question A when that question was asked first, and the probability of a 'yes' answer to question A when it was asked imediately after survey question B. We also calculate the three probability terms in our measure of dependence on the right-hand-side in Eq. (22): $P_*(B)$ (the probability of a 'yes' answer to survey question B when that question was asked first) $P_*(A)$ (the probability of a 'yes' answer to survey question A when that question was asked first) and $P_*(A|B)$ (the probability of a 'yes' answer to survey question A, given that a 'yes' answer to question B has just been given). Combining these terms as on the right-hand-side in Eq. (22) to compute the measure of dependence between questions for each survey, we find a reliable correlation between the order effect for questions and the measure of dependence ($r = 0.41$, $p = 0.0004$), supporting the noise model's prediction.

Following Wang et al. (2014), we also test this prediction by considering only those 17 surveys in this dataset where Wang et al.'s value Z (which measures the overall order effect across all combinations of questions in the survey; see Eq. (14)) had a value greater than 0.1. Wang et al. (2014) take this subset of surveys as showing the largest and most reliable effects of order on question answering. For this subset we again find a significant correlation between the order effect for questions and the dependence measure in Eq. (22) ($r = 0.71$, $p = 0.0001$), further supporting the noise model's prediction. Fig. 4 gives a scatterplot relating order effects for individual questions in these 17 surveys to values of their respective dependence measures. Taken together, these results suggest that the noise model does a reasonable job of predicting the magnitude and direction of order effects in sequential question answering, at least in cases where there is a significant order effect. In making this prediction, the noise model goes beyond the quantum probability account for order effects in sequential question answering.

5. General discussion

The main point of this paper has been to describe a systematic model which accounts for a series of apparent invariants or identities (the addition law, the Bayes rule identity, the QQ identity, and others) in people's probabilistic judgement. This model assumes that human probabilistic judgement is based on a rational process (one that follows frequentist probability theory) that is subject to random noise and error. Our secondary point here is to suggest that while an alternative model, based on quantum probability, can account for QQ identity results, it is unable to account for results on the various other identities.

In this section we address responses from quantum probability theorists to our work. We begin by considering some recent results from Yearsley and Trueblood (2017), who investigate the simultaneous occurrence of both sequential order effects and the conjunction fallacy in judgements about real-world events: the probabilities of various different Republican and Democratic candidates winning various U.S. presidential primaries. As part of this investigation, Yearsley & Trueblood test a specific prediction the noise

---

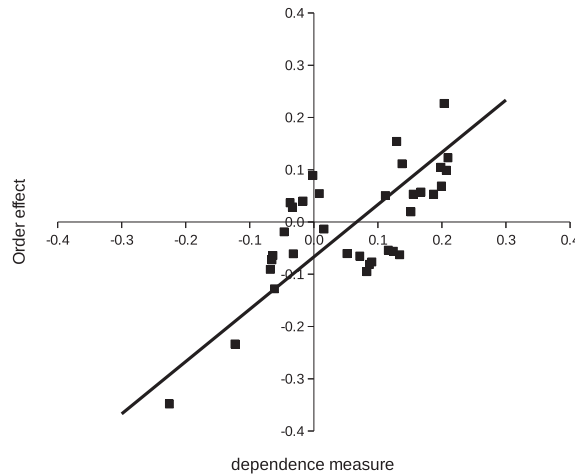[3] We would like to thank Zheng Joyce Wang for providing this data.

Fig. 4. Scatterplot of values of the order effects $P_{BA}(A) - P_e(A)$ and the associated dependence measure $P_e(B)[P_e(A|B) - P_e(A)]$, for the 17 surveys in Wang et al.'s dataset for which the order measure Z was greater than 0.1. Since each survey contained two questions, each with an associated order effect and dependence measure, there are 34 points in this scatterplot. There was a reliable correlation between order effect and dependence measure across these 34 points ($r = 0.713$, $p = 0.0001$). The line of best fit was calculated via Deming regression, which accounts for error in observations on both the x and y axes.

model makes about conjunction fallacy occurrence: that the conjunction fallacy should only occur in averaged probability estimates when the conjunctive probability has a value less than 0.5. This prediction follows from the role of regression in the noise model's account for the conjunction fallacy, which, in the model, 'pushes' conjunctive probabilities upwards when they are less than 0.5, and so can produce cases where the average conjunctive probability is greater than the average constituent probability: a conjunction fallacy in averaged probability estimates (see Costello & Watts, 2014, 2017, 2016b). Yearsley and Trueblood (2017) find occurrences of the conjunction fallacy in averaged estimates in cases where the conjunctive probability is greater than 0.5, apparently contradicting the noise model's prediction.

Our response to this is to point out that the noise model's account of the conjunction fallacy in terms of regression does not address order effects on probability estimation: that account assumes order-free presentation of probability questions, and describes the conjunction fallacy as arising solely due to regressive effects which push conjunctive probability estimates upwards. As described in the previous section, however in the noise model sequential order effects can push constituent probabilities downwards, especially if there is a negative dependence between events A and B. This downward movement of constituent probabilities can also produce the conjunction fallacy, by moving a constituent probability estimate below a conjunctive probability estimate. The results of Yearsley and Trueblood (2017), which find conjunction fallacy occurrence in cases where average conjunctive probabilities are relatively high and where there are reliable order effects, can thus been seen as consistent, at least to some degree, with the probability theory plus noise model: the model would explain these conjunction fallacy results as arising due to sequential order effects moving constituent probabilities downward, in addition to regressive effects moving conjunctive probabilities upwards. Support for this proposal relating order effects and the conjunction fallacy comes from the fact that there was a reliable correlation between conjunction fallacy rates and sequential order effects across individuals in Yearsley and Trueblood's data. Future work investigating this account in detail will clearly be necessary, however.

We now consider two more general responses from quantum theorists, both resting on the distinction we discussed earlier: the distinction, in quantum probability, between situations where people judge the probability of some event or statement (give a probability estimate, as in the addition law and Bayes rule tasks), and situations where people do not give a probability estimate but instead make a binary choice (give a yes–no answer to a question, as in the QQ identity tasks).

One line of response from quantum theorists here would be to argue that the Quantum Probability model is only designed to address binary choices, and not probability estimates. Since only the QQ identity involves binary choices, while the other identities involve probability estimates, this would mean that our evidence that the quantum model cannot account for results associated with those other identities does not count against that model. The problem here is that, as we described earlier, the quantum probability model has been presented in the literature both in terms of producing binary responses (e.g. Wang & Busemeyer, 2013; Wang et al., 2014) and in terms of producing probability estimates (e.g. Busemeyer et al., 2011; Trueblood & Busemeyer, 2011; Yearsley & Trueblood, 2017). If the quantum probability model does not, in fact, apply to probability estimation, but instead only applies to choice, the quantum model becomes unable to account for results on probability estimation: the scope of the model seems significantly narrowed.

A second line of response to our results is to extend or augment the quantum probability model by adding a regressive transformation operator that converts a person's 'internal' probability estimate (produced by state projection as in the equations of quantum probability theory) into an overt probability judgement response. If this transformation operator is chosen to be regressive in some way, then the overt probability judgements produced in the quantum model will be subject to regressive effects similar to

those produced by random error in the probability theory plus noise model, and such an augmented quantum probability model will be able to produce results consistent with the addition law and Bayes rule results (because these results are caused purely by regression in the noise model).

While this is certainly a feasible response to our argument against the quantum model here, it is problematic in a number of ways. First, such a transformation or extension of the quantum probability model will modify the model's predictions about probability judgement in general. It is thus not clear whether this new model will remain consistent with previous results supporting the quantum approach. Second, it is not clear whether there is any *a priori* reason for assuming a regressive transformation operator in the quantum probability approach, beyond simply fitting the data. The noise model, by contrast, provides an *a priori* reason for such regression: it arises as a consequence of purely random error. If there were a mechanism within quantum probability that produced regressive effects, that would provide significant support for the quantum approach. As far as we can see, however, quantum probability does not in itself give rise to any such regressive transformation. Finally, and most fundamentally, there is the problem of Occam's razor: 'entities must not be multiplied beyond necessity'. Our argument in this paper is that various biases in probability judgement and sequential question answering, and simultaneous patterns of agreement with identities such as the addition law, Bayes rule and the QQ identity, can be explained solely in terms of a single entity: regressive effects caused by random error. The quantum model, however, seems to require two entities to explain these results: a quantum model of 'internal' probability representation, and a regressive transform producing overt probability judgements. Given that the noise model seems to be able to account for these results via just one of these entities (that of regression), it is not at all clear what extra explanatory power the quantum model brings.

6. Conclusions

Much research on people's probabilistic reasoning over the last 50 years has focused on the various significant biases seen in probability estimation and judgement. Invariants such as the addition law, the Bayes rule identity, and the QQ identity, which hold simultaneously with these biases, reveal an important fact: they show us that these biases are systematically and quantitatively related and can be explained mathematically. We can see this in the case of the QQ identity, where there are reliable order effects (biases) in responses which nonetheless cancel out when responses are combined in the identity. We also see this in the addition law and Bayes rule identities, where there are reliable biases in probability estimates which again, cancel out when those estimates are combined in those identities.

We have shown that one particular mathematical model of probabilistic reasoning, the probability theory plus noise model, can account for all three of these invariants, while simultaneously explaining various systematic patterns of bias in probability estimation. By contrast the quantum probability model seems unable to explain agreement with the addition law and Bayes rule identities while simultaneously accounting for these patterns of bias. These results support the theoretical proposal in the probability theory plus noise model, which is that human probabilistic judgement is based on a rational process (one that follows frequentist probability theory) that is subject to random noise.

It is important to stress that we are not suggesting that people's probability estimates are themselves rational. This is clearly not the case: there is very extensive evidence demonstrating that people's probability estimates are systematically biased away from the requirements of probability theory. We argue that these biases are a consequence of the influence of random noise on the probability estimates generated by an underlying rational process. While this noise is random, it has systematic, directional effects (our noisy model's expected averages for probability estimates are systematically biased away from the 'true' probability values, in a way that seems to match the biases seen in people's estimates) which are cancelled out in these three identities. This model gives a new and useful perspective on the cognitive processes underlying people's probabilistic reasoning.

References

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

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

Costello, F. J., & Mathison, T. (2014). On fallacies and normative reasoning: when people's judgements follow probability theory. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.). *Proceedings of the 36th annual meeting of the cognitive science society* (pp. 361–366). .

Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review, 121*(3), 463–480.

Costello, F., & Watts, P. (2016a). People's conditional probability judgments follow probability theory (plus noise). *Cognitive Psychology, 89*, 106–133.

Costello, F., & Watts, P. (2016b). Probability theory plus noise: Replies to Crupi and Tentori (2015) and to Nilsson, Juslin and Winman (2015). *Psychological Review, 123*(1), 112–123.

Costello, F., & Watts, P. (2017). Explaining high conjunction fallacy rates: The probability theory plus noise account. *Journal of Behavioral Decision Making, 30*(2), 304–321.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). Minerva-DM: A memory processes model for judgments of likelihood. *Psychological Review, 106*(1), 180–209.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over- and underconfidence: The role of error in judgment processes. *Psychological Review, 101*(3), 519–527.

Fisher, C. R., & Wolfe, C. R. (2014). Are people naïve probability theorists? A further examination of the probability theory + variation model. *Journal of Behavioral Decision Making, 27*(5), 433–443.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: How noisy information processing can bias human decision making. *Psychological Bulletin, 138*(2), 211–237.

Kellen, D., Singmann, H., & Batchelder, W. H. (2017). Classic-probability accounts of mirrored (quantum-like) order effects in human judgments. In: *Decision.* https://doi.org/10.1037/dec0000080.

Moore, D. W. (2002). Measuring new types of question-order effects: Additive and subtractive. *The Public Opinion Quarterly, 66*(1), 80–91.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 41*(1), 1–20.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*(4), 273–286.

Trueblood, J. S., & Busemeyer, J. R. (2011). A quantum probability account of order effects in inference. *Cognitive Science, 35*(8), 1518–1552.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293–315.

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

Wang, Z., Solloway, T., Shiffrin, R. M., & Busemeyer, J. R. (2014). Context effects produced by question orders reveal quantum nature of human judgments. *Proceedings of the National Academy of Sciences, 111*(26), 9431–9436.

Yearsley, J. M., & Trueblood, J. S. (2017). A quantum theory account of order effects and conjunction fallacies in political judgments. *Psychonomic Bulletin & Review*. http://dx.doi.org/10.3758/s13423-017-1371-z.