## Title
Quantum Sequential Sampler: a dynamical model for human probability reasoning and judgments



**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

**Authors**
Huang, Jiaqi
Busemeyer, Jerome
Ebelt, Zo
et al.

**Publication Date**
2023

Peer reviewed

# Quantum Sequential Sampler: a dynamical model for human probability reasoning and judgments

**Jiaqi Huang (huangajq@iu.edu)[1], Jerome Busemeyer (jbusemey@indiana.edu)[1]**
**& Zo Ebelt (Zo.Ebelt@city.ac.uk)[2], & Emmanuel M. Pothos (emmanuel.pothos.1@city.ac.uk)[2]**
[1] Department of Cognitive Science, Indiana University, 1001 E. 10th Street, Bloomington, IN 47405 USA,
[2] Department of Psychology, City, University of London, 32-38 Whiskin Street, London, EC1R 0JD, United Kingdom

**Abstract**

Probability judgments appear to violate basic axioms of probability theory, which seems to contradict with the recent successes of Bayesian models of cognition. To explain these violations, we propose the Quantum Sequential Sampler model, which combines quantum probability for explaining conjunction and disjunction fallacies, and a sequential sampling model that maps subjective quantum probabilities into responses. Our model explains probability judgments by a dynamical process, and achieves state-of-the-art performance in the biggest dataset for probability judgments to-date. Comparing with existing Bayesian models, our model predicts both probability judgments and violations of probability identities better.

**Keywords:** quantum cognition; probabilistic reasoning; probability judgments; Bayesian cognition; sequential sampling; Markov process

## Introduction

Bayesian probability has had profound impacts in many areas of cognitive sciences, including perception (Knill & Richards, 1996; Kersten, Mamassian, & Yuille, 2004), memory (Hemmer & Steyvers, 2009; Schooler, Shiffrin, & Raaijmakers, 2001), intuitive physics (Sanborn, Mansinghka, & Griffiths, 2013; Xu et al., 2021), and causal reasoning (Tenenbaum & Griffiths, 2002; Hill, 2011). The optimality of Bayesian inference has also been supported from an evolutionary standpoint (Valone & Giraldeau, 1993; Luttbeg & Warner, 1999). On the contrary, there is also abundant empirical evidence showing apparent violations of Bayesian principles and Kolmogorov probability axioms in human probability judgments. The most famous examples are the conjunction fallacy (Tversky & Kahneman, 1983), disjunction effect (Tversky & Shafir, 1992), and the disjunction fallacy (Bar-Hillel & Neter, 1993). Given the successes of Bayesian models, understanding why they are in conflict with empirical evidence that probabilistic reasoning sometimes violates Bayesian principles, is perhaps one of the most important problems in cognitive sciences.

Many attempts have been made to resolve this conflict, and these attempts can be mainly divided into two categories (1) theories and heuristics that try to explain specific fallacies (2) quantitative models that predict the magnitude of probability judgments. Although theories that belong to the first category such as the inductive confirmation theory (Crupi, Fitelson, & Tentori, 2008; Tentori, Crupi, & Russo, 2013) provide valuable insights into particular fallacies, the focus of this paper will be on computational model in the second category.

A popular type of quantitative models are those that explain probability judgment fallacies through sampling biases and errors. These models treat probability judgment as a separate entity from subjective probability that follows Kolmogorov axioms, and construct a mapping from subjective probability to probability judgment through a noisy sampling process (Dasgupta, Schulz, Tenenbaum, & Gershman, 2020). Examples of these models include the probability plus noise model (Costello & Watts, 2017) and the Bayesian sampler model (Zhu, Sanborn, & Chater, 2020). Despite being successful in explaining various violations, these sampling-error models also arguably have some theoretical limitations. First, sampling-error models must rely on rounding mechanisms (Zhu et al., 2020). The reason is that these models generate predictions from a binomial distribution with small sample sizes, and thus without assuming people round numbers in a specific way, the likelihood of some responses given the model would have been zero. Indeed, there is good evidence that people round numbers in many inferential tasks (Budescu, Weinberg, & Wallsten, 1988; Wallsten, Budescu, & Zwick, 1993), but proposed mechanism for rounding are mostly task dependent (Ebelt, Pothos, Busemeyer, & Huang, 2022). On the other hand, a model that samples probability judgments from an approximately continuous distribution will be more preferable. Secondly, sampling-error models account for conjunction and disjunction fallacies through additional assumptions, such as that conjunctions and disjunctions are more likely to be subject to counting errors (Costello & Watts, 2017), or computationally more demanding to count than other probabilities (Zhu et al., 2020). There are alternative axiomatic accounts of these two fallacies such as the quantum probability models (Busemeyer, Pothos, Franco, & Trueblood, 2011; Busemeyer & Bruza, 2012; Pothos & Busemeyer, 2022). However, quantum probability models are also not the complete story of probability fallacies: because they are strictly axiomatized, they also tend to be less flexible in predictions and may fail to account for certain violations of probability identities (Costello & Watts, 2017).

To resolve these shortcomings, our approach is to combine quantum probability models' axiomatic explanation of conjunction and disjunction fallacies with an improved response model: instead of using biased frequencies from fixed sampling with small sample sizes, we employ a sequential sampling model as the response model. Sequential sampling



In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*. 

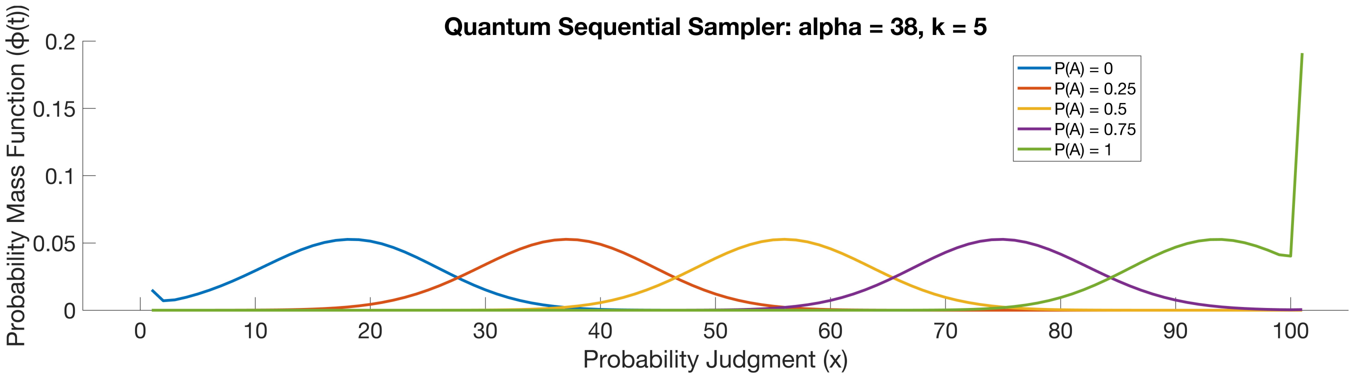


Figure 1: Illustration of the sequential sampling part of Quantum Sequential Sampler model. $P(A)$ stands for subjective probability of an arbitrary event $A$, $\alpha$ denotes the drift parameter, and $k$ denotes the additive bias parameter. Each curve represents the probability mass function over probability judgments developed from different subjective probabilities, under the same drift parameter $\alpha$ and additive bias parameter $k$.

models have several advantages over fixed sampling in modeling probability judgments: (1) the resulting response distribution can be continuous (2) sequential sampling model puts probability judgment into a dynamical picture. Sequential sampling models have already had a wide range of applications in cognitive sciences including decision making under uncertainty (Busemeyer & Townsend, 1993; Ratcliff & McKoon, 2008), categorization (Nosofsky & Palmeri, 1997), valued based decision making (Busemeyer & Townsend, 1993; Usher & McClelland, 2001), and memory recognition (Ratcliff & McKoon, 2008). However, to our knowledge, this is the first time sequential sampling models are applied to model direct probability judgments.

In the remainder of the paper, we will first explain the details of our novel model. We will then compare our model to the Bayesian Sampler model in predicting probability judgments and violation of probability identities in our new dataset[1] concerning probability judgments of the 2020 US presidential election. Finally, we will discuss some future directions inspired by our work.

## Quantum Sequential Sampler

The Quantum Sequential Sampler model **(QSS)** consists of two parts. The first part is a quantum probability model, and the second part is a sequential sampling model that maps the quantum probabilities into probability judgment responses. The approach of combining quantum probability with a sequential sampling model has also been applied to other fields (Rosendahl, Bizyaeva, & Cohen, 2020). Conceptually, the quantum probability part represents subjective probabilities establishes by participant's prior knowledge about the events, and the sequential sampling part represents the cognitive process of sequentially sampling evidence from mental simulations to estimate the subjective probability of these events. Comparing with models that estimate probability judgments through sampling with a fixed hypothetical sample size for computing frequencies, the sequential sampling approach instead assumes that people gather evidence from mental simulations for a fixed period of time. In the following, we will explain the mathematical details of our model.

### Quantum Probability Part

The quantum probability part is the model for conjunction fallacies in Busemeyer et al. (2011). The key reason why quantum theory can produce conjunction fallacies lies in the quantum interference term:

$$P(A \ \&\ then\ B) - P(B \ \&\ then\ A) = Int(AB), \tag{1}$$

where $P(A \ \&\ then\ B)$ is the quantum conjunctive probability of measuring arbitrary event $A$ first and then arbitrary event $B$, and $P(B \ \&\ then\ A)$ is the same computed from the reverse order. When $Int(AB) = 0$, we are essentially in the case of Bayesian probability, because all other probabilities except for conjunctions and disjunctions are computed in the same way in quantum probability as that in Bayesian probability (Busemeyer & Bruza, 2012). That said, the use of quantum probability for explaining probability fallacies is not necessarily in conflict with Bayesian approach of cognition for other tasks.

There are constraints of when conjunction and disjunction fallacies may occur, and it is important for QSS to satisfy these constraints to produce fallacies. For arbitrary events $A, B$ such that $P(A) \geq P(B)$, the quantum rules (Busemeyer & Bruza, 2012) suggest that a conjunction fallacy may only occur when

$$0 \leq P(B \ \&\ then\ A) \leq P(B) \leq P(A \ \&\ then\ B) \leq P(A), \tag{2}$$

that is only the conjunctive probability $P(A \ \&\ then\ B)$ may produce conjunction fallacy. And in this case, since $P(\neg A) \leq P(\neg B)$, we could also derive that a disjunction fallacy may

[1] Due to space limitations, we only describe this dataset in a rudimentary way here, so as to focus on the modeling details. A detailed presentation of the dataset will be in a manuscript by the present authors, which is currently in preparation.

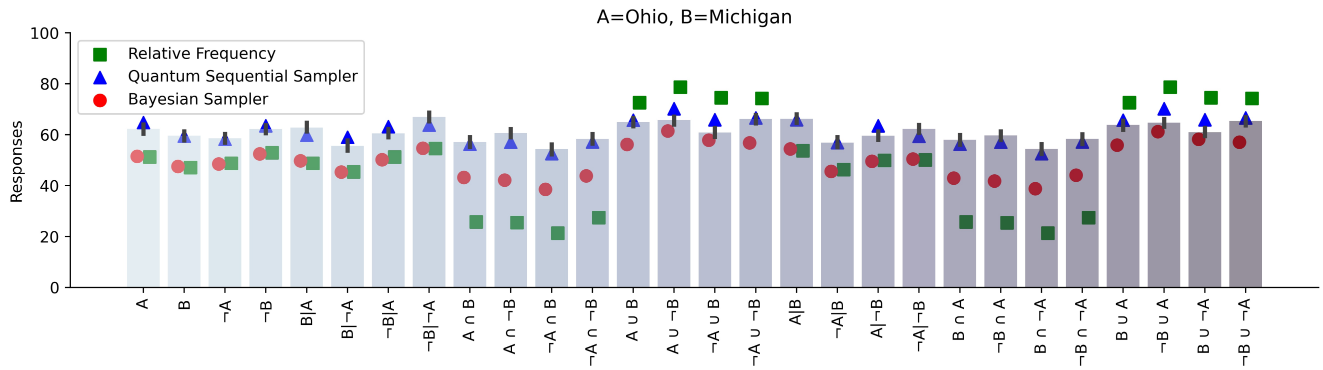


Figure 2: Mean predictions of the models over all participants compared with the data for each probability event in our novel dataset for probability judgments concerning 2020 election results, for states Ohio and Michigan. Events $A, B$ denote Biden will win the states, and events $\neg A, \neg B$ denote Trump will win the states. The bars show the distribution of the data with 95% confidence interval displayed by the error bar. [2]

only occur when

$$P(B) \leq P(B \text{ or then } A) \leq P(A) \leq (A \text{ or then } B) \leq 1. \quad (3)$$

More generally, Inequality 2 and 3 could also be shown to constrain the magnitude of the interference term $Int(AB)$ (Busemeyer et al., 2011).

In QSS, we employ what we referred to as the "more likely first" assumption: for arbitrary events $A, B$

$$P(A) \geq P(B) \implies P(A \wedge B) := P(A \;\&\; \text{then } B). \quad (4)$$

It is not hard to check that the more-likely-first assumption is indeed consistent with Inequality 2 for the conjunction fallacy and can be used to derive a disjunction fallacy.

## Sequential Sampling Part

A continuous-time Markov process is used as the sequential sampling part of QSS, because such processes have been widely supported as plausible mechanisms for how responses are produced from internal biases (Busemeyer & Townsend, 1993; Ashby & Waldron, 2000). When the Markov process is also continuous in space, the process becomes a drift diffusion process and generates continuous distribution (Busemeyer & Diederich, 2010).

To introduce our model, we start with the generic solution to the Kolmogorov forward equation with constant intensity:

$$\phi(t) = e^{Kt}\phi(0). \quad (5)$$

In the above, $K$ is the $N \times N$ intensity matrix that encodes the state transition rates, $\phi(0)$ is the initial distribution across the $N$ states, and $\phi(t)$ is what $\phi(0)$ will evolve into after time $t$. When probability judgment responses are measured as integers from 0 to 100, $N = 101$. To obtain $\phi(t)$, which encodes the likelihood distribution over probability judgments at response time $t$, we therefore need to define the intensity matrix $K$ and the initial distribution $\phi(0)$.

With a reflecting boundary condition, the general intensity matrix $K$, with the first index in the subscript representing row number and the second index representing the column number, can be written as:

$$\begin{aligned}
&K_{i,i+1} = \beta_+ \text{ for } 1 \leq i \leq N-1 \\
&K_{i+1,i} = \beta_- \text{ for } 1 \leq i \leq N-1 \\
&K_{i,i} = -(\beta_+ + \beta_-) \text{ for } 2 \leq i \leq N-1 \\
&K_{1,1} = -\beta_+ \\
&K_{N,N} = -\beta_-,
\end{aligned} \quad (6)$$

and $K_{i,j} = 0$, for any other indexes. For each $P(A)$, that is the quantum subjective probability of an arbitrary event $A$, where $A$ can be any of the isolated conjunct, conjunction, disjunction, and conditionals, we define

$$\begin{aligned}
\beta_+[P(A)] &= P(A) * \alpha + c_+ \\
\beta_-[P(A)] &= (1 - P(A)) * \alpha + c_-,
\end{aligned} \quad (7)$$

where $\alpha$ is the free parameter that describes the drift rate, and $c_+$ and $c_-$ are two additive biases defined by a single parameter $k$:

$$c_\pm = \begin{cases} 1 \text{ if } \pm k \leq 0 \\ \pm k + 1 \text{ if } \pm k > 0. \end{cases} \quad (8)$$

In the above, $k$ defines whether people underestimate or overestimate their probability judgments: when $k > 0$, the Markov process has a bias to drift towards the boundary of probability judgment of 100, and thus generates overestimation of judgments; vice versa, when $k < 0$, the model generates underestimation. The reason for focusing on k as opposed to the

[2] All judgments appear to be above 50% in Figure 2 because of an overestimation bias presented empirically. Similar overestimations were also found for unpacking effects (Sloman, Rottenstreich, Wisniewski, Hadjichristidis, & Fox, 2004) and other uncertainty measures (Epping & Busemeyer, 2023). We double-check our methods, and they are correct.

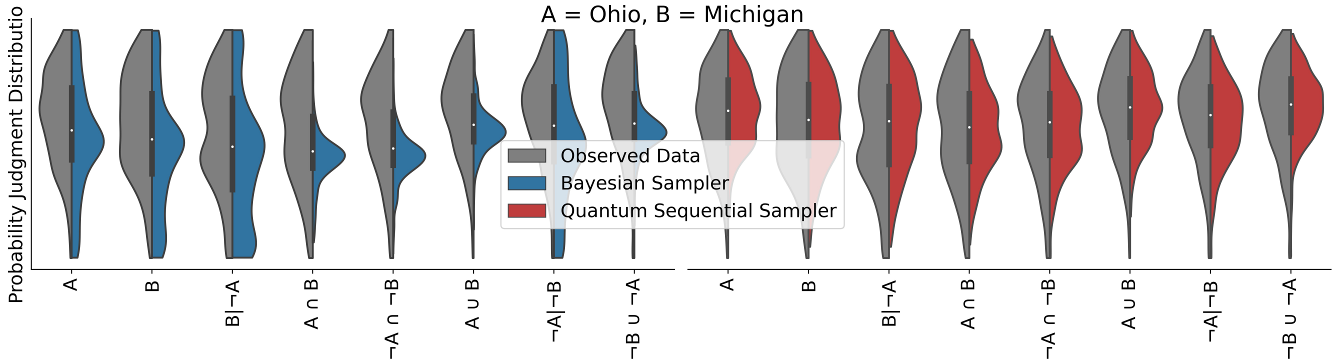


Figure 3: Violin plots showing the distribution of predictions of the models compared to the observed data for the Ohio and Michigan pair. $A, B$ denote Biden will win the states, and $\neg A, \neg B$ denote Trump will win the states.

unit drift rate is that $\beta_+, \beta_-$ must always be positive by definition, and only the difference between $\beta_+$ and $\beta_-$ matters for the direction of the drift in a Markov process. Conceptually, $c_\pm$ describes preexisting judgment biases of the evidence accumulation process, which are then regulated by a mental simulation process described by the $P(A)$ part.

For all $P(A)$, we assume that the initial distribution is always the symmetric Beta distribution $Beta(a,a)$ discretized into an $N \times 1$ vector. The use of Beta distribution is consistent with previous work (Zhu et al., 2020; Dasgupta et al., 2020) in modeling probability inferences and judgments. The specific discretization method we employ is the following: let $\psi(x)$ be the probability density function of $Beta(a,a)$; then the initial distribution of the Markov process is defined as:

$$\begin{aligned}
&\phi(0)_i \propto \psi(\frac{i-1}{100}) \text{ for } 2 \leq i \leq 100 \\
&\phi(0)_1 \propto \psi(0.005) \\
&\phi(0)_{101} \propto \psi(0.995), \qquad (9)
\end{aligned}$$

where the subscript represents the column indexes of $\phi(0)$. $\phi(0)$ will then be normalized according to the above definition.

Figure 1 illustrates the final distributions $\phi(t)$ over probability judgments, for various subjective probabilities. In general, the further away the subjective probability is from 0.5, the faster $\phi(t)$ drifts towards the reflecting boundaries and reaches the stationary distribution. With a positive additive bias parameter $k$, the final distributions drift towards the right-hand-side boundary faster than it drifts towards the left.

## Bayesian Sampler

We compared our model with the Bayesian Sampler (BS) model (Zhu et al., 2020). BS assumes that previous experience establishes a symmetric Beta prior distribution $Beta(\beta,\beta)$, which is updated with a mental sampling process that encodes the subjective probability. Formally, for arbitrary event $A$, let $N(A)$ be the sample size of mental sampling for $A$ and $P(A)$ be the subjective probability, and let $S(A) \sim Bin(N(A), P(A))$ and $F(A) = N(A) - S(A)$ count the number of instances $A$ occurs and does not occur in the mental sampling correspondingly. Zhu et al. (2020) assumed that participants report the beta posterior means as their responses:

$$R_{BS}(A) = \frac{S(A)+\beta}{N(A)+2\beta}. \qquad (10)$$

$R_{BS}(A)$ will then be a binomial random variable, as $S(A)$ follows a binomial distribution.

## Model Comparison

### Dataset

The dataset for model comparison, to be reported fully in our forthcoming work, involves 1162 participants and 78 responses per participant, which is the biggest dataset of probability judgments to-date. The probability judgments concern whether Biden or Trump will win particular states in the 2020 US election, where Biden winning is considered as the complement of Trump winning, assuming the chance of other candidates winning close to zero (the study was conducted after the democratic primary but before the election). The 78 responses include all marginal, conditional, conjunctive, and disjunctive probability judgments of the election outcomes of three pairs of states. Conjunctive and disjunctive probability judgments were measured with two different orders of the conjuncts to test for potential order effects. Without going into details, we note that conjunction and disjunction fallacies and violations of probability identities were found for the majority of participants.

### Model Fitting

**Quantum Sequential Sampler** QSS was fitted by setting $k$ and $\alpha$ to be the same across all events, which is also the simplest version of the model. Conceptually, in this version of the model, the additive biases $c_\pm$ represent general over and under estimation biases across all events.

For any probability judgment response $R(A)$, the likelihood of $R(A)$ to be generated by the Quantum Sequential Sampler (QSS) with response time $t$ is given by:

$$L(R(A),t|QSS) = \phi(t)_{R(A)} = (e^{K[P(A)]t}\phi(0))_{R(A)}. \qquad (11)$$

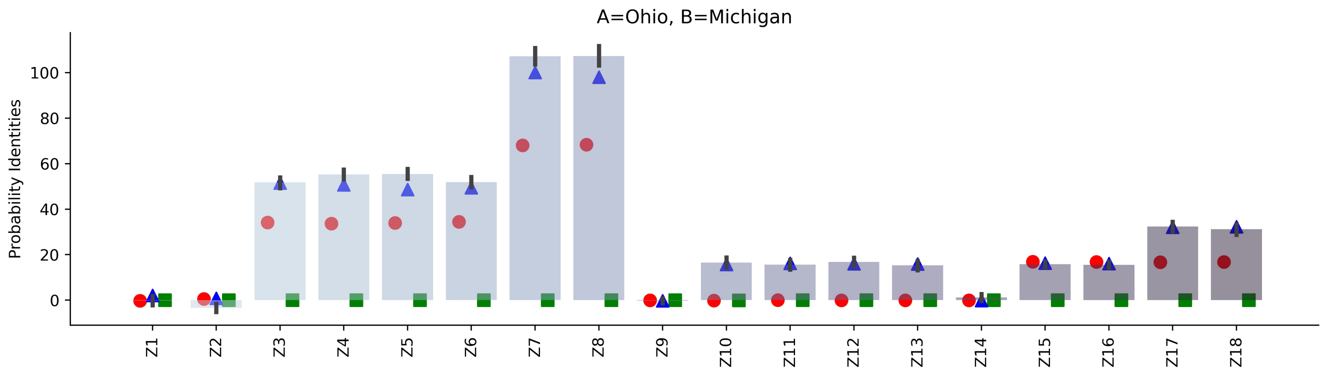


Figure 4: Predictions of the Z identities in Zhu et al. (2020), for the Ohio and Michigan pair. The bar plot shows the Z identities computed from the mean data with 95% confidence interval displayed by the error bars, and the shapes show the Z identities computed from the corresponding models' predictions.

Since response time is not measured, we make the simplest assumption that $t$ is the same across all events and thus can be absorbed into other parameters.

**Bayesian Sampler** One problem with the version of BS in Zhu et al. (2020) is that probability judgments follow a binomial distribution that may not include all integers from 0 to 100. To circumvent this issue, Zhu et al. (2020) assumed that people rounded responses from their predictions. However, the rounding bias of reporting 5s and 10s in verbal responses reported by Zhu et al. (2020) was observed to a lesser extent when responses are measured in rating scales in our new dataset. Thus, to fit BS to our new dataset, we need to modify BS as it is in Zhu et al. (2020) to assume that responses are directly sampled from the posteriors rather than being the posterior means. Our modified approach is consistent with many other previous works in Bayesian cognition (Tenenbaum, Griffiths, & Kemp, 2006; Griffiths, Kemp, & Tenenbaum, 2008), where the posterior distribution is directly used for judgments. The likelihood of response $R(A)$ to be generated by BS for this modified approach is computed as:

$$L(\frac{R(A)}{100}|BS) = \sum_{x=0}^{N(A)} B_n(x) * B(\frac{R(A)}{100}), \tag{12}$$

where $B_n(x)$ is the probability mass function of the binomial distribution $Bin(N(A), P(A))$, and $B(x)$ is the discretized probability density function of the Beta posterior $Beta(\beta + x, \beta + N(A) - x)$. Note that $B(x)$ is discretized in the same way as that in Equation 9 for the initial distribution of QSS. Finally, Zhu et al. (2020) assume that $N(A)$ is the same for all marginals and conditionals, but could be smaller for conjunctions and disjunctions, as a way to generate fallacies. We therefore fit two sample size parameters for the Bayesian Sampler model, as consistent with the original work.

## Fitting Comparison

QSS was compared with BS through the Bayesian Information Criterion (BIC), with the results summarized in Table 1. QSS has a much lower mean BIC than BS, and both models are much better than the baseline model, that only makes uniformly random guesses. Between all of the participants, we found that QSS outcompetes BS for 66% (769) of participants, which agrees with the mean BIC result.

We also compared the models using generalization tests. In particular, we performed the generalization tests in two ways: (1) fit the models on all other probabilities except for the conjunctions and test on predicting the conjunctions (2) the same except now test on predicting the disjunctions. The mean $G^2$ across all participants for both conditions are shown in Table 1. QSS performs the best overall, but BS outperforms QSS slightly when testing on disjunctions. Between participants, QSS outcompetes BS for 78% of participants when testing on conjunctions and 52% when testing on disjunctions.

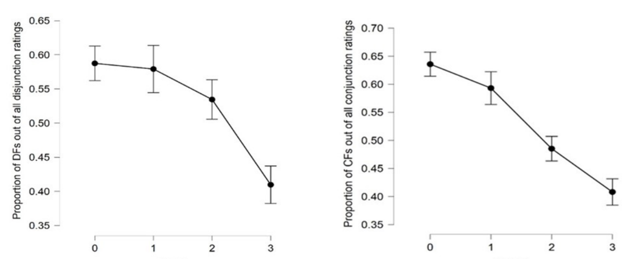


Figure 5: The proportion of conjunction (CF) and disjunction fallacies (DF) as a function of the scores of the cognitive reflection tests (CRT). Participants who make more intuitive judgments have a lower CRT score, and vice versa for participants who make more analytical judgments.

## Predictions

Besides comparing the two models through BIC, we also examined the predictions of the models. The predictions of QSS

| Model | mean BIC | Conj test | Disj test | k |
|---|---|---|---|---|
| QSS | 609.42 | 183.84 | 209.63 | 10 |
| BS | 662.94 | 231.81 | 205.61 | 9 |
| Uniform | 718.41 | 221.05 | 221.05 | 0 |

Table 1: The BIC and generalization tests results. A model (Uniform) which uniformly randomly guesses integer from 0 to 100 was also fitted. Mean BIC is the average BIC score for the model over all the participants, and k is the number of parameters for the model. Conj test is the mean $G^2$ when the models were tested on the conjunctions and were fitted on everything else, and Disj test is the same when the models were tested on the disjunctions by a generalization test. For both BIC and $G^2$, the lower the score, the better the model.

were computed as the expected value of probability judgment given the final distribution $\phi(t)$, while the prediction of BS was the mean of the binomial beta distribution (Zhu et al., 2020). We also compared the two models' prediction to that of a relative frequency model [3] which is expected to follow the Kolmogorov axioms.

We computed the mean predictions over all participants for each model and the result for a single pair of states is shown in Figure 2. QSS predicts the conjunction and disjunction probabilities much better than BS, while BS also shows a better prediction overall compared to the relative frequency model.

Besides plotting the mean predictions, we also constructed violin plots for the distributions of model predictions against the data across all participants in Figure 3. According to the Figure, QSS also captures the distributional characteristics of judgments better than BS for our dataset.

### Probability Identities

Zhu et al. (2020) listed 18 probability identities that must be zero according to Kolmogorov axioms, but are often violated empirically. We compare QSS with BS in predicting these identity violations for our new dataset. The result is shown in Figure 4. QSS outperforms BS in predicting most of the violations of identities. Compared to the dataset in Zhu et al. (2020), the present dataset shows a much larger magnitude of violations for most of the identities. This may partially account for why the Bayesian Sampler predicted the probability identities worse than it did in Zhu et al. (2020).

### Quantum Interference

Finally, although our model is named the Quantum Sequential Sampler model, the subjective probabilities do not have to be quantum: when all quantum interference terms are zero, QSS is Bayesian. We ran a $G^2$ test for each participant and fitted one version of the model with quantum interference and the other simpler version using only Bayesian probability to check whether the additional quantum interference parameter is significant. We found 403 participants out of 1162 (35%) show a significant quantum interference effect. Despite observing that several participants can be explained by Bayesian subjective probability, there are still a considerable amount of participants who show significant quantum effects, which raises the interesting question: when will participants be quantum, and when will they be completely Bayesian? One possibility could be that participants who make more intuitive judgments are more likely to be quantum than participants who make more analytical judgments. Intuitive and analytical participants can be distinguished using a cognitive reflection test (CRT) (Toplak, West, & Stanovich, 2011). Our study found that the proportion of conjunction and disjunction fallacy is correlated with the CRT scores of the participants: the more intuitive the participants, the more likely they commit to the fallacies (see Figure 5). Similar correlations were also observed in previous research (Trueblood, Yearsley, & Pothos, 2017).

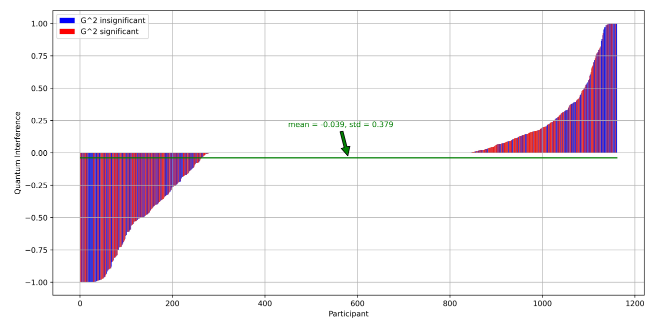


Figure 6: Ranked quantum interference parameters for all of the participants. The green horizontal line shows the mean and the standard deviation of the parameters. The parameters that are statistically significant are colored in red.

## General Discussion

We proposed a novel Quantum Sequential Sampler model for probability judgments. The model takes advantage of both the axiomatic explanation of conjunction and disjunction fallacies from quantum theory, and the power of the sequential sampling framework for translating subjective probabilities into responses. Our model achieves superior performance in fitting our new dataset over the Bayesian Sampler model (Zhu et al., 2020).

Our model opens many future inquiries for understanding probability in cognition. First, the sequential sampling part of our model allows us to jointly predict response time and probability judgments. In the present work, since no response time data was collected in the present dataset, we cannot test any predictions concerning response times yet. Future work can check whether the joint predictions of our model on probability judgments and response times are correct. Second, as mentioned, future work may also explore the conditions when participants are more likely to show significant quantum interference effect.

[3] Since the likelihood function of the relative frequency model is binomial, the model was fitted by sum of square error rather than BIC.

## Acknowledgment

EMP was supported by European Office of Aerospace Research and Development (EOARD) grant FA8655-23-1-7220. JRB was supported by the Air Force Office of Scientific Research under award number FA9550-20-1-0027.

## References

Ashby, F. G., & Waldron, E. M. (2000). The neuropsychological bases of category learning. *Current Directions in Psychological Science*, *9*(1), 10–14.

Bar-Hillel, M., & Neter, E. (1993). How alike is it versus how likely is it: A disjunction fallacy in probability judgments. *Journal of Personality and Social Psychology*, *65*(6), 1119.

Budescu, D. V., Weinberg, S., & Wallsten, T. S. (1988). Decisions based on numerically and verbally expressed uncertainties. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 281.

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

Busemeyer, J. R., & Diederich, A. (2010). *Cognitive modeling*. Sage.

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.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review*, *100*(3), 432.

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.

Crupi, V., Fitelson, B., & Tentori, K. (2008). Probability, confirmation, and the conjunction fallacy. *Thinking & Reasoning*, *14*(2), 182–199.

Dasgupta, I., Schulz, E., Tenenbaum, J. B., & Gershman, S. J. (2020). A theory of learning to infer. *Psychological review*, *127*(3), 412.

Ebelt, Z., Pothos, E., Busemeyer, J., & Huang, J. (2022). A large-scale study into human judgments: new evidence for probabilistic fallacies. *British Psychological Society Bulletin*.

Epping, G. P., & Busemeyer, J. R. (2023). Using diverging predictions from classical and quantum models to dissociate between categorization systems. *Journal of Mathematical Psychology*, *112*, 102738.

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition.

Hemmer, P., & Steyvers, M. (2009). A bayesian account of reconstructive memory. *Topics in Cognitive Science*, *1*(1), 189–202.

Hill, J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, *20*(1), 217–240.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as bayesian inference. *Annu. Rev. Psychol.*, *55*, 271–304.

Knill, D. C., & Richards, W. (1996). *Perception as bayesian inference*. Cambridge University Press.

Luttbeg, B., & Warner, R. R. (1999). Reproductive decision-making by female peacock wrasses: flexible versus fixed behavioral rules in variable environments. *Behavioral Ecology*, *10*(6), 666–674.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological review*, *104*(2), 266.

Pothos, E. M., & Busemeyer, J. R. (2022). Quantum cognition. *Annual review of psychology*, *73*, 749–778.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural computation*, *20*(4), 873–922.

Rosendahl, L., Bizyaeva, A. S., & Cohen, J. (2020). A novel quantum approach to the dynamics of decision making. In *Cogsci.*

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological review*, *120*(2), 411.

Schooler, L. J., Shiffrin, R. M., & Raaijmakers, J. G. (2001). A bayesian model for implicit effects in perceptual identification. *Psychological Review*, *108*(1), 257.

Sloman, S., Rottenstreich, Y., Wisniewski, E., Hadjichristidis, C., & Fox, C. R. (2004). Typical versus atypical unpacking and superadditive probability judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 573.

Tenenbaum, J., & Griffiths, T. (2002). Theory-based causal inference. *Advances in neural information processing systems*, *15*.

Tenenbaum, J., Griffiths, T. L., & Kemp, C. (2006). Theory-based bayesian models of inductive learning and reasoning. *Trends in cognitive sciences*, *10*(7), 309–318.

Tentori, K., Crupi, V., & Russo, S. (2013). On the determinants of the conjunction fallacy: probability versus inductive confirmation. *Journal of Experimental Psychology: General*, *142*(1), 235.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275–1289.

Trueblood, J. S., Yearsley, J. M., & Pothos, E. M. (2017). A quantum probability framework for human probabilistic inference. *Journal of Experimental Psychology: General*, *146*(9), 1307.

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

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological science*, *3*(5), 305–310.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological review*, *108*(3), 550.

Valone, T. J., & Giraldeau, L.-A. (1993). Patch estimation by group foragers: what information is used? *Animal Behaviour*, *45*(4), 721–728.

Wallsten, T. S., Budescu, D. V., & Zwick, R. (1993). Comparing the calibration and coherence of numerical and verbal probability judgments. *Management Science*, *39*(2), 176–190.

Xu, K., Srivastava, A., Gutfreund, D., Sosa, F., Ullman, T., Tenenbaum, J., & Sutton, C. (2021). A bayesian-symbolic approach to reasoning and learning in intuitive physics. *Advances in Neural Information Processing Systems*, *34*, 2478–2490.

Zhu, J.-Q., Sanborn, A. N., & Chater, N. (2020). The bayesian sampler: Generic bayesian inference causes incoherence in human probability judgments. *Psychological review*, *127*(5), 719.