Towards a Bayesian Model of Data Visualization Cognition

Yifan Wu* UC Berkeley
Larry Xu† UC Berkeley
Remco Chang ‡ Tufts University
Eugene Wu § Columbia University

Abstract

Data visualizations are often used to assist decision making with probabilistic data. Different cognitive biases can affect the accuracy of user insights gained during the visual analytics process. However, evaluating bias in visualization usage is challenging and difficult to quantify. In this paper, we propose a Bayesian inference model based on cognitive science research to fill this gap. We outline the details for this model and the evaluation steps, including an end-to-end demonstration experiment that we performed. The results provide initial validation for using a Bayesian inference model to quantitatively measure bias in visual analytics.

1 Introduction

Data visualization is increasingly an integral part of data-driven decision making, and much research effort has been put into understanding aspects of cognition around visual data analytics (e.g., [5, 11, 21, 22, 27] to name a few). However, despite the effort, cognitive bias remains difficult to measure and detect quantitatively. Past studies have focused on graphical perceptions, which have proven critical to guide the design of effective visualizations [6]. In contrast, with few exceptions (e.g. [26]), less work has been done on how visualization designs influence users' beliefs.

Intuitively, detecting such influence is simple. As shown by Pandey et al. [26], to detect a user's change in belief after viewing a visualization, one would measure people's belief prior to and after seeing the visualization. This structure of measuring beliefs "prior to" and "after" viewing a visualization strikes a similarity to the concept of Bayesian statistics where "priors" and "posteriors" are used to measure the probability of the occurrence of an event. As such, a potential mechanism to detecting cognitive bias can be viewed as a measure of how much a visualization changes a person's beliefs beyond the Bayesian posterior.

The advantage of this approach is that one can quantify changes in a person's belief state. In the context of visualization, this means that we can model how and to what degree a visualization changes a user's beliefs, and create a shared basis to compare across visualization designs. For example, when evaluating how a visualization changes a user's mind, the difference between the expected Bayesian posterior of a user's belief and the measured user's belief could quantify the bias induced by the visualization design.

Researchers in cognitive science have been studying whether human cognition could be modeled with Bayesian statistics and found that a probabilistic framework holds the promise of a complete theory of learning—able to predict how beliefs change with observations [10, 12, 13]. Further more, many "irrational" behaviors could be modeled as extensions to the Bayesian model [1, 20, 28], so not only could cognitive biases be detected, but also modeled in a principled way.

*e-mail: yifanwu@berkeley.edu
†e-mail: larry.xu@berkeley.edu
‡e-mail:remco@cs.tufts.edu
§e-mail: ew2493@columbia.edu

In this paper, we demonstrate the use of Bayesian inference to model and measure the change of people's beliefs as the degree by which it deviates from Bayesian posteriors. We discuss how to set priors for participants, compute the theoretical Bayesian posterior, and elicit the posterior from participants. We use a running visual analytics example to explain the set up of a Bayesian model, and a demonstration experiment that we piloted to illustrate how the model can be applied to quantitatively measure changes in beliefs. Lastly, we discuss ways to extend the basic Bayesian inference model introduced in this paper with the rich literature from cognitive science to capture different cognitive biases.

This work shows potential promise for a more robust way to model and quantify cognition of visual analytics. We hope to inspire discussion on Bayesian inference modeling methods for cognitive bias studies.

2 Background

Increasingly, diverse evaluations of the cognition of visualizations are being done by the information visualization community, ranging from type and amount of insights generated [24], how memorable different visualization designs are [3], how trustworthy sampled data seem [7], statistical perception and understanding [4, 15], and data recall/comprehension [18]. We present a brief overview of their approaches to motivate the need for a general, quantifiable model, and why Bayesian inference is a good candidate.

Liu et al. investigated interaction in information visualization as actively constructing and manipulating mental models, for three primary purposes: external anchoring, information foraging, and cognitive offloading [22]. In a similar spirit, we want to explore ways to model finer grained statistical understanding of the data being visualized.

Pandey et al. compared the persuasive power of charts versus tables [26]. The authors described different ways people are persuaded and not persuaded, suggesting that it is difficult to isolate the effect of the medium and the effect of the topic, and that persuasion has many facets and succinct quantitative representations are difficult. This motivates our work to seek a way to address these challenges.

Kim et al. found that visualization of the difference between the actual outcome and people's expectations can enhance recall, and hints at the modeling of the process as participants making inferences on existing priors [18].

Fisher et al. evaluated the effect of sampling on user trust, suggesting that direct visual representation of uncertainty, e.g. error bars, could inspire enough trust for the system to be useful [7]. This suggests that users are incorporating probabilities into their reasoning process. Zhao et al. modeled interaction in visual analysis as hypothesis testing, and provide tools to prevent spurious discoveries caused by the presence of multiple hypotheses [32]. Some have argued for a Bayesian analysis approach as opposed to the traditional hypothesis testing method for its flexibility and ability to incorporate new information cumulatively [19]. More recently, Kangasrääsiö et al. have propose using approximate Bayesian computation to parametrize cognitive model from behavioral data [17].

From these recent research, a trend of probabilistic modeling of the visual analysis process is emerging. We find illuminating research from cognitive science and propose a candidate model: Bayesian inference. Bayesian reasoning is not new to information visualization—researchers have investigated whether visualizations help improve human Bayesian reasoning, [4, 23, 25]. Not only is Bayesian reasoning challenging for people, research have found even understanding uncertainty is challenging [15].

Despite these issues, modeling human cognition with Bayesian models has been pioneered by cognitive science researchers [10, 12, 13]. Modeling the analysis as Bayesian updates does not require explicit human Bayesian reasoning—instead of giving participants equations to calculate, they are expected by to answer based on *intuition*.

Griffiths and Tenenbaum performed experiments that revealed for everyday activities and found that humans are surprisingly consistent with real world statistics and the Bayesian model [14], contrary to previously well established theories by Tversky and Kahneman [30]. More broadly, they have been able to successfully use the model to explain previous findings of modeling word learning, property induction, and causal learning, and irrationality, such as seeking confirmation and anchoring [13]. Other work by Bayesian modelers have also suggested ways to elicit priors from people [29], which makes the modeling more practical.

These emerging theories and tools for modeling has, to the best of our knowledge, not yet been applied to information visualization and visual analytics. Much of current research suggest that modeling the visual data analysis process is ripe for a Bayesian investigation. The goal of this paper is to explore whether it is possible to use Bayesian inference to model the visual analytic process, in the hope of providing a quantitative and expressive model for understanding and measuring cognitive bias for visual data analytics.

3 Bayesian Modeling of Beliefs

Bayes' law describes the probability of an event, based on previous knowledge of the event and data points seen. It is stated as

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where A and B are two events, and $P(B|A)$ is the likelihood of B happening conditioning on A being true. $P(B)$ can be evaluated by marginalizing over alternative events to A, denoted A', without loss of generality, as the cases where A did not happen: $P(B) = P(B|A)P(A) + P(B|A')P(A')$. This yields the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')} \tag{1}$$

While the above equation seem far removed from visualizations, it is possible to map the viewing of data visualization as a Bayesian inference process: treat B as the event of seeing a visualization of certain data, and A as an outcome which the visualization user has some belief over. Then setting $B =$ viewing a visualization of a set of data and $A =$ belief of outcome in Equation (1) derives an analytical solution to the Bayesian posterior.

Concretely, we introduce a running example for the rest of the paper, illustration the modeling, experiment design, and result analysis. Take a fairly common business scenario where people evaluate whether a company is doing well based on viewing the visualization of the company's sales data. In the real world, the considerations may be complex, but for the sake of the experiment, we define two kinds of companies: one with a 60% probability of sales increasing month to month (hereafter referred to as "strong"), and another with 40% ("weak").

Based on the sales data, we could formulate an analytical solution to the Bayesian posterior using Equation (1) by supposing $A =$ company is strong, $A' =$ company is weak, and $B =$ visualization(sales data). P(company is strong) and P(company is weak), shown as $P(A)$ and $P(A')$ in Equation 1, are the known as the "priors". These two values represent the viewer's prior belief that a company would be strong or weak respectively. Determining the priors is considered one of the key challenges in using Bayesian models. We will discuss methods for setting (or eliciting) a viewer's prior beliefs in Section 3.1 and demonstrate in Section 4.

The conditional probability of seeing a visualization of a sequence of sales numbers for a strong company, $P(\text{visualization(sales data)}|\text{company is strong})$ (which is $P(B|A)$ in Equation (1)) can be calculated based on the sales data and the likelihood of increase/decrease month to month. Since the sequence is assumed to be independent, the probability follows the product rule for independent events (if C and D are independent, $P(C,D) = P(C)P(D)$).

For instance, if the sales data for a type of tea is, for the months January to June is $10, 21, 29, 19, 35, 42$, then the probability of seeing this sequence, purely based on what we know about strong company's ability to improve month to month sale, is $p \cdot p \cdot (1-p) \cdot p \cdot p$, since the changes were increase, increase, decrease, increase, and increase, which simplifies to $p^4(1-p)$ (4 increases and 1 decrease).

To summarize, $P(\text{visualization(sales data)}|\text{company is strong})$ can be computed as follows:

$$\begin{aligned} P(B|A) = & P(\text{increase}|\text{company is strong})^{(\text{number of increase})} \\ & \cdot (1 - P(\text{increase}|\text{company is strong}))^{(\text{number of decrease})} \end{aligned} \tag{2}$$

The same formulation can be used to compute the conditional probability of $P(\text{visualization(sales data)}|\text{company is weak})$ (which is $P(B|A')$ in Equation (1)). Together with $P(A)$ and $P(A')$, we can compute $P(B)$ (the denominator of Equation (1)).

3.1 Measuring Priors

As noted earlier, one potential challenge in using Bayesian statistics is accurately eliciting priors (i.e. determining the values for $P(A)$ and $P(A')$). There has been recent research that explores technique to elicit prior knowledge. Kim et al. proposed elicitation by explicit self-explanation in text format and predicting the data before seeing the data [18]. While these methods are worth experimenting, the problem of elicitation may be avoided by directly setting the priors and informing the participants of them. The challenge of this approach is that the participants might not be able to fully contextualize and comprehend the meaning of these priors. While it is generally known that people are inherently bad at understanding and calculating probabilities, important work in psychology have found that the right representation, frequency format, could significantly improve human comprehension of probabilities [8]. Using this knowledge, we inform the users of the prior by describing the frequencies of strong companies versus weak companies, as such *50 out of every 100 companies are strong companies*.

For visualizations scenarios with real world data and existing beliefs, it is not always to set the priors. In this case, the experiment could ask the users directly. For instance, *how many companies currently in the market do you think qualifies as "strong companies" per our definition out of 100*?

In cases where it is difficult for the people to describe their priors explicitly, there are other more involved research methods that could elicit the priors passively. Transmission chains are one technique to extract priors [2, 29]. By passing information from one person to another, the original information in the input gets erased and the latent biases surface, creating a stable end result that represents the priors.

3.2 Sources of Biases

If we assume that Bayesian statistics could model human cognition of data visualizations, then the differences between the analytical results from the Bayesian model (e.g. the computed value of $P(A|B)$) and the user responses are due to biases. These biases could be either introduced by the visualization, for which different visualizations will have different effect on users' responses, or inherent to the users' thinking process, cognitive abilities, or previous experiences etc., which should remain unchanged regardless of the visualizations.

In this section we describe a few ways to extend the basic Bayesian setup to potentially more accurately model user responses. We will describe three different models as examples, but to prescribe situations for which they are applicable, if at all, requires future studies. This section is to inspire discussion and future research ideas.

The first is **prospect theory** by Kahneman et al, which models the phenomena that people are risk averse when sure gains are present, and risk seeking when sure losses are present [16]. This could be applied to cases when the visual analysis involves decisions that might yield gain or losses. For instance, correctly identifying the company to be strong could have investment return consequences.

The analytical form of incorporating prospect theory is found in [9], where the linear model is

$$w(p) = a + (b - a) \cdot p$$

and curved model

$$w(p) = \frac{\sigma \cdot p^{\gamma}}{\sigma \cdot p^{\gamma} + (1-p)^{\gamma}}$$

We could apply the weight function w on the probability reported. That is to say, if the participant reported that the company is 0.9 likely to be strong, the actual belief p' is calculated by solving $a + (a-b) \cdot p' = 0.9$ (by the linear model). The parameters a, b (in the linear model), or σ, γ (in the curved model) could be computed by fitting against the experiment data.

Second is when participants do not trust the data, either due to data source validity, or how representative the data is, both of which were suggested in [26] as factors which people have considered when evaluating whether a visualization is persuasive. This could be incorporated to the model by **discounting the observations**. For example, instead of observing the original number of increase (in Equation 2), the participants observe effectively a discounted amount, say a discount factor of θ of the original. This modifies Equation (2). Concretely for example, the number of month-to-month increase becomes (number of increase $\cdot\, \theta$). θ could by found by searching through different values of $0 < \theta < 1$ for lowest error.

Third is when the participants have existing biases towards a topic—i.e. **inherent priors**, which is also reported in [26] as "anchoring to core beliefs". In the running example, when we inform the subjects the companies are 50% likely to be strong, it might translate to 40% likely to be strong, if, for instance, the subject has just read a news article about economic decline.

These three types of adjustment to the model are just examples extensions to the Bayesian model, many more could be explored, which is one advantage of the Bayesian model: the flexibility despite simple formulation [13].

4 Demonstration Experiment

For a concrete end-to-end example, we present a between-subject experiment that evaluates how two visualizations inform users differently, one static and one interactive (as shown in Figures 1 and 2 respectively)[1]. We chose to compare the two visualizations because while the interactive visualization may help people get a more intuitive sense of the changes, which has proven to have some positive effect on understanding statistics [15], the static visualization showing summary data as a line chart is more common, which could make the task of estimating probabilities easier.

[1]The experiment could be viewed at `https://www.gpaas.xyz/gpaas/experiment/seabass/seabass_pilot/?workerId=demo` If the experiment does not progress, try changing the "workerId" to a random string in the URL

Using the example of determining the strength of a company discussed in Section 3, our experimental procedures are:

1. We inform the participants of the priors, e.g. "For this task, 50 out of every 100 companies are strong companies", and show a corresponding visualization that draws rectangles whose width correspond to the percentage of strong companies, as shown in the Figure 3.

2. The participants then view/interact with the visualization, either the static one (Figure 1), or the interactive one (Figure 2). In both cases, the participants see the month-to-month sales of a fictitious tea company.

3. We ask the participants to predict the probability of this tea company being a strong company. The participants would enter their responses by dragging a slider valued from 0 to 100 (see Figure 3). The default value of the slider is set at the prior given (i.e. 50%). Additionally we asked the participants to report their confidence of their response on a 3-point Likert scale.

The data is randomly generated but with the constraint that the month-to-month sales change is within a small range so that one randomly large increase does not cause unintentional bias. All sales data are within 0 and 100 (with unit unspecified), and with no correlation between the three products.

Since the experimental setting involves the mention of multiple probabilities, we recognize that the instructions can be difficult to follow for some participants. We therefore included three training tasks in the experiment where we gave the participants feedback for what the expected Bayesian posteriors were (without explaining the mathematical calculations). Further studies are needed to ensure a more robust and accurate prescription of the likelihood and prior in similar experiments. One potential method is to provide the participant with more examples of what it means to have the likelihood and the probabilities, which may not be too far removed from real life when experts gain experience via past events.

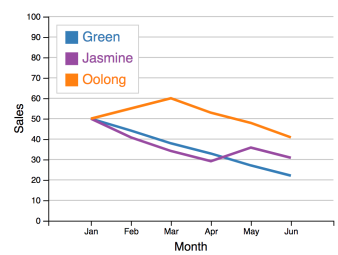

Figure 1: Static line chart visualization of each month's result, each line represents the product, with data over six months and 15 data points total of month to month changes.

5 Demonstration Evaluation

In this section we describe the methods for analyzing the participants' responses to quantitatively measure how much change a visualization causes in a participants' beliefs. Note that the goal of this section is primarily to illustrate how to use the model and show evidence for its promise, and not focused on the actual implications on visualizations, which requires more experimentation and analysis.

The analytical solution to the Bayes equation 1, i.e. the Bayesian posterior, provides the unbiased baseline of what we expect the true probability to be given the data. When the Bayesian posterior is compared to the participants' responses, the difference of the two quantifies the bias. The comparison can be done in a variety of

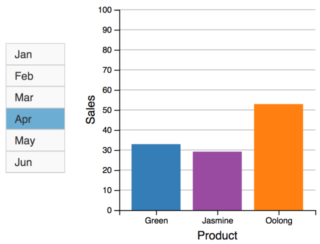

Figure 2: Interactive bar chart visualization of each month's result, where as the subject hovers over the month the corresponding sales information shows up. The subject could get an intuitive sense of change by the bar height changes.

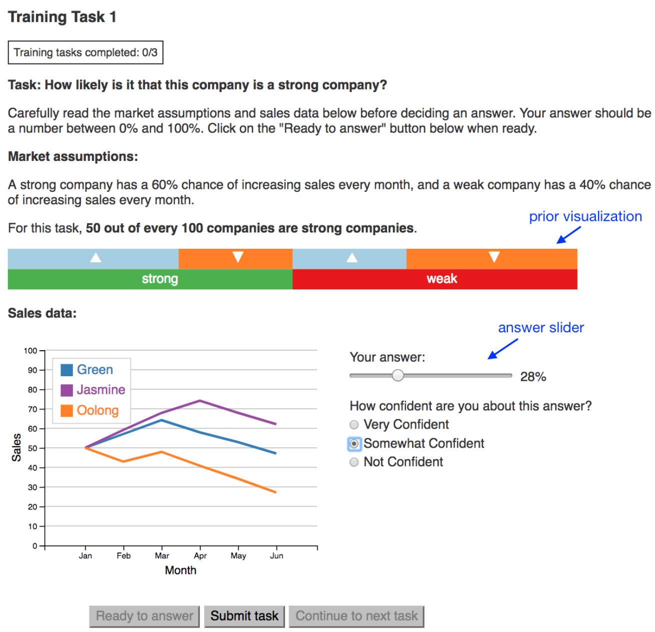

Figure 3: Experiment design overview, including the simple visualization of the prior and the fixed conditional probability information, and the response collection interface.

ways, but one simple method is to compute the average distance of the participants' responses to the Bayesian posterior. Besides capturing the absolute signed bias, we could also aggregate the signed bias, which emulates practices in prediction markets where people's beliefs are pooled [31]. Since the difference is signed, we could gain insights into the *direction* of the bias—is the posterior higher, or lower than expected.

In addition, correlating the participants' responses and the Bayesian posterior using Pearson's r could evaluate how much of the variance of the participant responses could be explained by the Bayesian model—the higher the correlation, the more effective the model is at describing and predicting the participants' responses. The correlation could be used to evaluate the effectiveness of different extensions discussed in Section 3.2.

For the pilot demonstration experiment, we found that while individual responses had varying degrees of error, on average 20%. The Bayesian model explains about 60% of the variance (by Pearson's r, with p value of less than $1e$-10), which suggests that the Bayesian model is appropriate for this experiment. The average signed error was close 0%, which is a rather surprising result, but consistent with the effectiveness of the crowd in prediction market experiments and other Bayesian cognition modeling experiments [12]. However again the experiment is more for demonstration purposes and the findings require more experimentation to be conclusive.

Using the bias measures, we could compare across different visualizations. In this pilot study, we found no significant difference between the static and interactive visualization ($p > 0.4, U = 11620$). However there was significant difference in terms of completion time ($p < 0.005, U = 9273$), by only about 20%, where the total time interactive visualization is 12.7 seconds, the median for static is 10 seconds.

One concern about the experiment is that people might not actually understand the priors we describe in the experiment. The results show that conditioning on the same data types (characterized by the total number of month-to-month increases), responses with the same prior are different $p < 0.1$ from responses with a different prior—this means that people must have used the prior to reason, because otherwise the distributions would be indistinguishable, since other factors are kept the same.

Figure 4 illustrates this point—the box plots in three different colors, denoting different priors 0.1 for red, 0.5 for green and 0.1 for blue, are distributed differently given the same data observed. It can be seen that given the same x-axis value (holding the amount of evidence the same), that participants' posteriors differ somewhat consistently [2] with the bayesian theoretical posterior, illustrated by the filled lines in Figure 4.

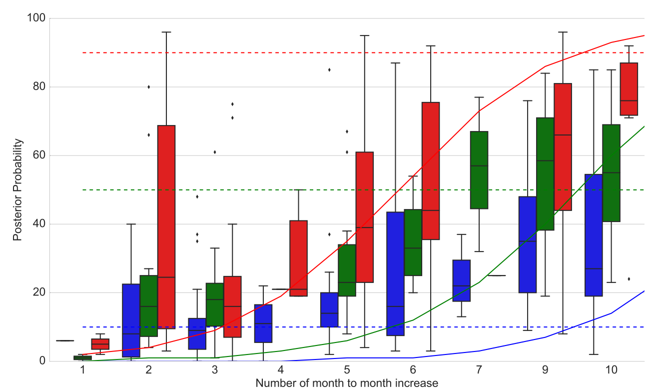

Figure 4: A visualization of difference by the prior of the probability of the company being "strong", red is 0.9, green is 0.5, and blue is 0.1. The dashed lines show priors, the filled lines show the theoretical Bayesian posteriors against the varying data observed (keeping likelihoods fixed), and the box plot of user responses, in comparison to the filled lines of the theoretical posterior.

6 Conclusion and Future Work

In this paper, we proposed to model cognition of information visualization as Bayesian inference through a description of the model and demonstration experiment. First, we presented the Bayesian model for visual analytics and walked through an example of how the model can be used to study the change in beliefs when using two different visualizations. We then demonstrated the procedures for a demonstration experiment and how to analyze the results.

The demonstration experiment shows one way to make use of the Bayesian model to evaluate the differences between a static visualization and an interactive one for how they shape the viewer's opinion about the underlying products.

[2]The noise/wide range between quartiles, especially around 2 and 7, may be due to the small number of experiment subjects.

Similar experiments could be designed to evaluate the inferences in experiments in Kim et al.'s recent work on visualizing the gap of one's prediction and reality by evaluating the posterior of the category of data as opposed to recall [18], or to put a more quantitate evaluation to Pandey et al.'s work on the persuasive power of data visualization [26]. However more research is needed to apply the Bayesian framework to less abstract settings with complex stimuli.

We hope that this work could inspire more quantitative evaluations of cognitive bias in information visualization and modeling of human cognition of information visualization.

Acknowledgments

We thank Falk Lieder at the Computational Cognitive Science lab at UC Berkeley for his helpful discussions on experiment design, and suggestions for how to model the results.

This research was partially supported by the following grants: NSF 1527765, NSF 1564049, NSF 1452977, and DARPA FA8750-17-2-0107.

References

[1] J. L. Austerweil and T. L. Griffiths. Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, 35(3):499–526, 2011.

[2] F. C. Bartlett and C. Burt. Remembering: A study in experimental and social psychology. *British Journal of Educational Psychology*, 3(2):187–192, 1933.

[3] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.

[4] G. L. Brase. Pictorial representations in statistical reasoning. *Applied Cognitive Psychology*, 23(3):369–381, 2009.

[5] R. Chang, C. Ziemkiewicz, T. M. Green, and W. Ribarsky. Defining insight for visual analytics. *IEEE Computer Graphics and Applications*, 29(2):14–17, 2009.

[6] W. S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association*, 79(387):531–554, 1984.

[7] D. Fisher, I. Popov, S. Drucker, et al. Trust me, i'm partially right: incremental visualization lets analysts explore large datasets faster. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1673–1682. ACM, 2012.

[8] G. Gigerenzer and U. Hoffrage. How to improve bayesian reasoning without instruction: Frequency formats. *Psychological review*, 102(4):684, 1995.

[9] R. Gonzalez and G. Wu. On the shape of the probability weighting function. *Cognitive psychology*, 38(1):129–166, 1999.

[10] N. D. Goodman, J. B. Tenenbaum, J. Feldman, and T. L. Griffiths. A rational analysis of rule-based concept learning. *Cognitive science*, 32(1):108–154, 2008.

[11] T. M. Green, W. Ribarsky, and B. Fisher. Building and applying a human cognition model for visual analytics. *Information visualization*, 8(1):1–13, 2009.

[12] T. L. Griffiths. Statistics and the bayesian mind. *Significance*, 3(3):130–133, 2006.

[13] T. L. Griffiths, N. Chater, C. Kemp, A. Perfors, and J. B. Tenenbaum. Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in cognitive sciences*, 14(8):357–364, 2010.

[14] T. L. Griffiths and J. B. Tenenbaum. Optimal predictions in everyday cognition. *Psychological science*, 17(9):767–773, 2006.

[15] J. Hullman, P. Resnick, and E. Adar. Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering. *PloS one*, 10(11):e0142444, 2015.

[16] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, pp. 263–291, 1979.

[17] A. Kangasrääsiö, K. Athukorala, A. Howes, J. Corander, S. Kaski, and A. Oulasvirta. Inferring cognitive models from data using approximate bayesian computation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1295–1306. ACM, 2017.

[18] Y.-S. Kim, K. Reinecke, and J. Hullman. Explaining the gap: Visualizing one's predictions improves recall and comprehension of data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1375–1386. ACM, 2017.

[19] J. K. Kruschke. What to believe: Bayesian methods for data analysis. *Trends in cognitive sciences*, 14(7):293–300, 2010.

[20] F. Lieder, T. L. Griffiths, Q. Quentin, and N. D. Goodman. Empirical evidence for resource-rational anchoring and adjustment. *Psychonomic Bulletin and Review*, 2017.

[21] Z. Liu, N. Nersessian, and J. Stasko. Distributed cognition as a theoretical framework for information visualization. *IEEE transactions on visualization and computer graphics*, 14(6), 2008.

[22] Z. Liu and J. Stasko. Mental models, visual reasoning and interaction in information visualization: A top-down perspective. *IEEE transactions on visualization and computer graphics*, 16(6):999–1008, 2010.

[23] L. Micallef, P. Dragicevic, and J.-D. Fekete. Assessing the effect of visualizations on bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2536–2545, 2012.

[24] A. V. Moere, M. Tomitsch, C. Wimmer, B. Christoph, and T. Grechenig. Evaluating the effect of style in information visualization. *IEEE transactions on visualization and computer graphics*, 18(12):2739–2748, 2012.

[25] A. Ottley, E. M. Peck, L. T. Harrison, D. Afergan, C. Ziemkiewicz, H. A. Taylor, P. K. Han, and R. Chang. Improving bayesian reasoning: the effects of phrasing, visualization, and spatial ability. *IEEE transactions on visualization and computer graphics*, 22(1):529–538, 2016.

[26] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The persuasive power of data visualization. *IEEE transactions on visualization and computer graphics*, 20(12):2211–2220, 2014.

[27] M. Pohl, M. Smuc, and E. Mayr. The user puzzleexplaining the interaction with visual analytics systems. *IEEE transactions on visualization and computer graphics*, 18(12):2908–2916, 2012.

[28] A. Shenhav, S. Musslick, F. Lieder, W. Kool, T. L. Griffiths, J. D. Cohen, and M. M. Botvinick. Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, (0), 2017.

[29] J. W. Suchow, M. D. Pacer, and T. L. Griffiths. Design from zeroth principles. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 2016.

[30] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making*, pp. 141–162. Springer, 1975.

[31] J. Wolfers and E. Zitzewitz. Prediction markets. *The Journal of Economic Perspectives*, 18(2):107–126, 2004.

[32] Z. Zhao, L. De Stefani, E. Zgraggen, C. Binnig, E. Upfal, and T. Kraska. Controlling false discoveries during interactive data exploration. *arXiv preprint arXiv:1612.01040*, 2016.