Fumeng Yang | Research Statement

> My research is situated at the intersection of human-computer interaction and visual computing, drawing methodology from statistics and artificial intelligence (AI). I aim to empower users in decision-making processes driven by model predictions.

Predictive models are essential tools for forecasting the future, now deeply integrated into our society. From real estate companies basing their strategies on price predictors [1] to voters influenced by election forecasts [2, 3], their impact is vast. However, all models can be flawed and are always uncertain, and human interpretation of them often falls short. These combined imperfections lead to poor decisions with significant consequences, ranging from financial losses in the millions of dollars [1] to widespread controversies, such as those seen after the 2016 U.S. presidential election [4].

I employ visualizations to help users overcome the imperfections inherent in both predictive models and humans, thereby steering them towards appropriate decisions. I create, deploy, and assess visual explanations for ❶ flawed or ❷ uncertain predictions to calibrate user trust and inform the decision-making process. I also utilize *vision science* and *virtual reality* to address known human limitations in interpreting visualizations, focusing on ❸ perceptual precision and ❹ memory capacity.

A defining characteristic of my research is its systematic interdisciplinary approach. I draw upon theory and methodology from various fields as the foundation for my experimental design, and craft data analyses grounded in deep knowledge of the domain in question. I have collaborated closely with over 30 scholars and students across a broad spectrum of disciplines, including *Bayesian statistics*, *computer vision*, *explainable AI*, *journalism*, *political science*, *perceptual psychology*, and *UI/UX design*.

My endeavors have resulted in notable publications at top-tier conferences and journals, including ACM CHI, ACM IUI, IEEE VIS, IEEE TVCG, and UbiComp (Fig. A). I have received **one best paper award** [2] and **three honorable mentions** [5, 4, 6], serving as **a first author on each**. Looking ahead, my future research will encompass other advanced computational techniques, tailor them for diverse users, and extend to include model design. My ultimate goal is to enable individuals and organizations to harness the full advantages of computational technology in decision-making.

A. Research Overview

My work tackles **four areas of imperfection** in **predictive models** and **human interpretations** to steer users towards appropriate decisions.

❶ **flawed prediction**
calibrating trust
Yang et al. IUI '20 ★ (paper award)

❷ **uncertain prediction**
conveying uncertainty
Yang et al. CHI '23a ★
Yang et al. VIS '23 ★
Yang et al. under review

❸ **perceptual precision**
quantifying precision
Harrison et al. VIS '14
Yang et al. TVCG '18
Ondov et al. VIS '20
Yang et al. CHI '23b

❹ **memory capacity**
enhancing capacity
Yang et al. TVCG '20
Yang et al. VIS '20 ★
Yang et al. TVCG '22

Model Explanations to Foster Trust in Decision-making

Trust is a crucial factor in users' decisions about whether to follow a model prediction [5]. Users may undervalue a model, leading to *undertrust*, or they may overestimate model reliability, leading to *overtrust* (Fig. B). Both undertrust and overtrust can result in inappropriate decisions that embrace mistakes or miss opportunities [1]. Thus, **appropriate decisions** require users' **appropriate trust** in predictive models.

However, machine learning models are often nontransparent and complex. It is nearly impossible for end users with limited expertise to anticipate ❶ a flawed prediction and calibrate their trust accordingly. This necessitates digestible and informative model explanations, which well-designed visualizations can provide. In this pursuit, I and my collaborators explored visual explanations for a support vector machine classifier. These explanations use training instances to illustrate how the classifier arrives at a prediction, enabling the identification of a prediction error. Our evaluation experiment asked participants to classify leaf instances, and demonstrated that all our visual explanations reduce overtrust and undertrust, thereby fostering appropriate trust. Our instantiation of human-machine collaboration yielded almost all appropriate decisions that followed correct recommendations and rejected incorrect ones. This work received **a best paper honorable mention award at ACM IUI** as exemplary research that conceptualizes appropriate trust [5].

B. Definition of appropriate trust

Users' appropriate trust in predictive models is defined by both prediction correctness and their decisions.

user decision \ model prediction	correct	incorrect
follow	Appropriate Trust	Overtrust
not follow	Undertrust	Appropriate Trust

Predictive models are also 2 **inherently uncertain**. For end users to trust models appropriately and make appropriate decisions, they must be **informed of the uncertainty in predictions**. Communicating such uncertainty to the general public has become a critical challenge as well as a prominent issue in the fields of *visualization*, *journalism*, and *statistics*. In recent years, U.S. election forecast visualizations produced by media outlets, such as FiveThirtyEight's 2020 presidential forecast [7], have attracted millions of viewers. The general public's misinterpretation of these visualizations has led to controversies surrounding election outcomes, which further influence voters' perception of democracy and their participation in elections [2]. In response to these challenges, since the beginning of my postdoc, I have collaborated with scholars and students from *journalism*, *political science*, and *perceptual psychology*, and conducted a series of studies on election forecast visualizations.

To explore the problem space, I first conducted formative studies. I surveyed 134 voters from the 2016 and 2022 U.S. presidential elections, and constructed more than 40 visualization prototypes. Following this, I conducted qualitative interviews with 13 participants from the survey, obtaining design lessons for forecast visualizations. Concurrently, my postdoc mentor, a Ph.D. student in our lab, and I developed a cognitive debiasing technique for improving forecast communication. This technique adjusts a forecast distribution based on a model of subjective probability, and it effectively reduces intrinsic bias in people's probability interpretation [4]. This work received **a best paper honorable mention award at ACM CHI** for its innovative and effective approach to uncertainty communication, as well as its rigorous methodology.

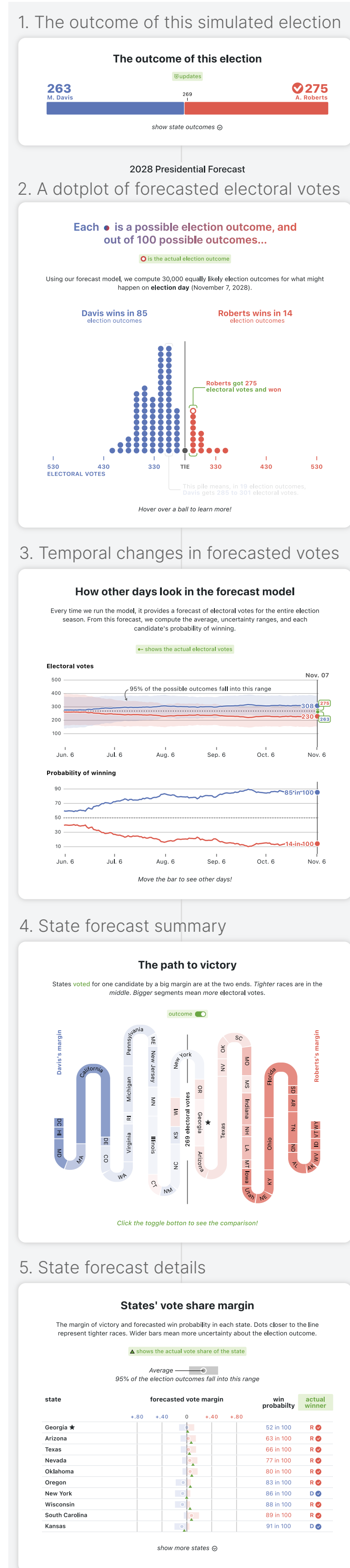
Guided by these preliminaries, I led **a longitudinal experiment during the 2022 U.S. midterm elections** to gauge the real-world impacts of forecast visualizations [2]. I coordinated with 9 co-authors, built a probabilistic forecast model for the governor elections, and constructed a forecasting website. The experiment ran from October 18, 2022 to November 23, 2022 and collected survey responses from 1,327 participants in 15 states. I then computed demographically representative estimates using Bayesian multilevel modeling and post-stratification on the U.S. 2020 Census data. I found that among the four visualizations tested, two (e.g., dotplot in Fig. C2) strongly heightened people's emotions and slightly influenced voting intention. They also gained people's trust the most but exacerbated affective polarization between the two major political parties. This work received **a best paper award at IEEE VIS**, with the Best Papers Committee stating, "The study is researched, planned and executed in an exemplary manner. The experimental results offer intriguing insights into the effects of these uncertainty visualizations on viewer emotions, trust, and intention to vote. Beyond the topical subject, the findings are useful to any uncertainty visualization endeavors."

I then turned to **high-profile, high-stakes U.S. presidential election forecasts** [3]. Trust in these forecasts influences public confidence in democratic processes, and thus, I explored techniques that can uphold this trust. I simulated different win probabilities for a hypothetical presidential election, created a professional-looking website interface (Fig. C), and conducted three experiments with varied uncertainty displays in an incentivized voting task. My Bayesian sequential decision-making models showed that text summaries and dotplots fostered the highest levels of trust in election forecasts over time; both partisan differences and educational backgrounds substantially influenced viewers' preferences and trust. Concluding from this work, I formulated recommendations for presenting U.S. presidential forecasts that can sustain the general public's trust over time and bridge viewers from diverse educational backgrounds.

I plan to conduct a study during the 2024 U.S. presidential election, which will incorporate live election results, a project that a Ph.D. student in our lab has been advancing under my mentorship. I anticipate making significant steps in conveying uncertainty to the general public and guiding them towards appropriate decisions.

C. Snapshot of our forecast website

This figure shows a simulated U.S. presidential election. Each panel below illustrates one facet of the forecast model or the outcome.



Vision Science and Virtual Reality for Model Communication

Despite my efforts to create visualizations for predictive models, **humans have limited capacities to receive and process information with precision**. To ensure that people receive sufficient and accurate information for decision-making, I also aim to improve perceptual precision and enhance human information processing capabilities. This involves delving into *vision science* to quantify **perceptual precision**, as well as exploring *cognitive science* and *virtual reality* to expand **memory capacity**.

Quantifying ③ perceptual precision lays the groundwork for creating visualizations that minimize perceptual errors. In collaboration with experts in *perceptual psychology*, I have worked on correlation perception in visualizations for nearly ten years. Correlation between two variables is a fundamental statistical concept. This research sheds light on how users perceive summary statistics from visualizations.

My initial work in 2013 [8] modeled the precision of correlation perception across various visualizations using Just-noticeable Difference (JND). This concept refers to the minimum difference required to detect a change in perception, such as noticing a difference in the area of two circles half the time, a threshold known as a JND (Fig. D). This alignment between low-level perception and statistical summaries sparked our speculation that people might rely on visualization features, such as bounding boxes, when judging correlation (Fig. E). By extracting a set of features to predict judgments in scatterplots, I found that four features were more predictive of people's judgments than correlation values themselves, suggesting that these features might be used in their judgments [9]. This speculation was later expanded to include mean and range judgments in bar charts [10]. Recently, in collaboration with experts in *computer vision* and *explainable AI*, I trained twenty-nine convolutional neural network (CNN) architectures on image datasets of correlation judgments [11]. The results revealed that VGG networks produce generalizable predictions across various scatterplots and help identify previously unrecognized features. Collectively, this body of my research provides a theoretical framework for understanding visualization perception and offers practical insights for minimizing perceptual errors through feature adjustments.

Visual ④ memory capacity is critical for many tasks [6]. To explore its role in visualization interpretation, I collaborated with scholars in *perceptual* and *cognitive sciences*, and conducted an experiment in which participants were tasked with recalling data from visualizations [6]. This research challenged the established design guideline that ranks the effectiveness of visual channels based on people's performance in a two-value ratio judgment task (Fig. F). After analyzing the experimental data, I discovered this guideline faltered with increased data points, as constrained by visual memory capacity. Instead, I proposed new, preliminary guidelines that adapt according to the number of data points (Fig. G). This work received **a best paper honorable mention award at IEEE VIS**, recognized for its evidence-based challenge to convention and its potential to reshape standards.

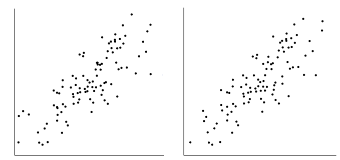
To enhance memory capacity, one approach is to use spatialization and mnemonic devices (e.g., a memory palace), which are effectively supported in virtual reality (VR). I conducted two experiments to explore the potential of VR for model communication. First, I developed a VR memory palace that enabled users to read scholarly articles and associate them with spatial cues in an immersive environment [12]. This technique improved recall performance, surpassing both the non-mnemonic baseline and a desktop-based memory palace, providing the first evidence of VR's support for reading-intensive tasks. Second, I explored the use of VR to communicate model performance [13]. I extracted hidden layer outputs from CNNs, employed dimension reduction techniques, and rendered 2D and 3D scatterplots of these outputs in VR and on a desktop. The results indicate that VR can match desktop effectiveness in conveying model performance.

D. Area perception 

The difference in the two circles might be noticeable only half the time.

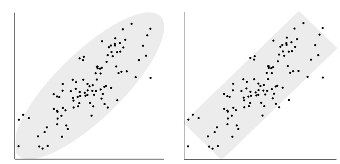
Correlation perception

The difference in correlation coefficients of the two datasets is just noticeable.



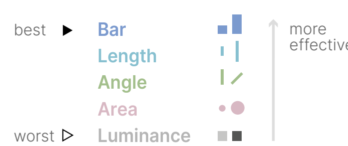
E. Visualization feature examples

People may use a prediction ellipse or a bounding box in judging correlation.



F. Conventional ranking of channels

The conventional ranking is based on a two-value ratio judgment task, guiding designers in choosing the most effective channels for data encoding.



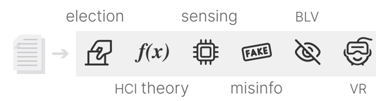
G. Our ranking based on a recall task

Each column ranks visual channels for a different number of data points (2 to 8). The channels on the top are the least biased and thus the best (▶).

Future Research Agenda

As highlighted in this statement, my research is characterized by its systematic interdisciplinary approach, encompassing a range of scientific domains. These research experiences have broadened my perspective and motivated me to explore ways that empower individuals and organizations to harness the advantages of computation. Drawing continual inspiration from my past research, I have maintained a document to track my evolving ideas (Fig. H). To date, this document has accumulated 38 potential projects across more than 10 topics, including other advanced computational techniques, visualization tools, pose sensing [14, 15], public policy, accessibility, perception and cognition, and more. Below, I select a few that particularly excite me.

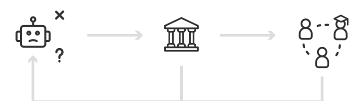
H. Example topics in my idea doc



❶ Communicating predictive models for policymaking. Predictive models are integrated into decision-making in high-stakes fields like law and business. My past work on election forecasting [2, 3] primarily targeted laypeople and has only begun to delve into these disciplines for domain experts and policymakers. The decisions they make can impact millions of people and may lead to significant societal consequences if marred by mistrust or errors in models (Fig. I). A notable example is the overtrust in AI, which often functions as a predictive model. I plan to bring my visualization approach to public policymaking, striving to calibrate trust and enhance fairness and reliability in decision-making processes. Additionally, my past research has revealed significant differences in trust and preference across diverse educational backgrounds [3, 5], motivating me to examine the roles of literacy [16] and numeracy in trust and decision-making. Besides my current ties in *political science* and *journalism*, I will also look for collaborative opportunities in *economics*, *law*, and *computer science education*.

I. Predictive models in policymaking

In policymaking, any mistrust or model error could generate significant societal consequences. I plan to use visualizations to help prevent such cases.



Also, the educational backgrounds of domain experts and the general public may significantly impact their trust and actions. I plan to explore these areas.

❷ Developing and democratizing visualization tools for advanced analytics. Advanced analytics encompass both predictive and prescriptive analytics. Prescriptive analytics succeeds predictive analytics by suggesting the best possible actions for optimizing practices (Fig. J). I completed a graduate-level course in prescriptive analytics and was fascinated by its use of computational techniques to provide precise recommendations (e.g., to reduce manufacturing costs), noting its impact on business strategy and operational processes. I have observed that visualization is extremely helpful for debugging and trusting an optimization process, yet surprisingly underutilized in this discipline. Recognizing the opportunities, I plan to work with domain experts to integrate visualizations with prescriptive analytics, streamlining the process.

J. What is prescriptive analytics

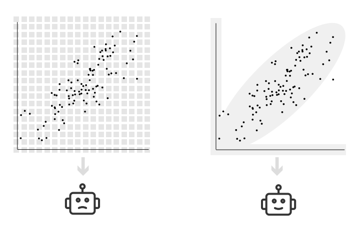
descriptive analytics	what happened?
predictive analytics	what will happen?
prescriptive analytics	what should we do?

Constructing these tools, however, demands high proficiency in visualization, design, and programming. To democratize this process, I will explore automating data-driven visual designs. Unlike current methods that produce static content, my approach will involve designing algorithms (e.g., combining them with generative AI) to decompose a draft or an existing visualization and to generate an editable code template automatically. I will seek collaborative opportunities in the field of *programming languages*.

❸ Designing domain-specific models for visual computing. Accurate predictions of human perception in visualizations can be used to prevent misinterpretation, guide design processes, and support decision-making. My prior work used existing deep learning techniques for predicting perceptual judgments [11], with limited accuracy and generalizability. To enhance both, I speculate that models used for visualizations will differ from those for natural images, such as CNNs. To this end, I plan to design model architectures specifically for visualizations (Fig. K). I am currently informally advising two ongoing projects led by Ph.D. students: Ouxun Jiang (Northwestern University), who is quantifying human perception of animation, and Sungbok Shin (University of Maryland, College Park), who is integrating human perception into machines. Alongside my existing ties with *vision science*, *explainable AI*, and *computer vision*, I will seek new collaborators in these fields to explore innovative directions.

K. Models for image vs. visualization

Models designed for natural images usually operate at a pixel level (left). Visualization images may operate on perceptual features (right). I plan to design models based on these concepts.



Icons used in my statements are either designed by me or licensed under Flaticon.

References

The IEEE Visualization and Visual Analytics (VIS) Conference is in liaison with IEEE Transactions on Visualization and Computer Graphics.

[1] Zillow's algorithm-fueled buying spree doomed Its home-flipping experiment. https://www.bloomberg.com/news/articles/2021-11-08/zillow-z-home-flipping-experiment-doomed-by-tech-algorithms.

[2] **Fumeng Yang**, Mandi Cai, Chloe Mortenson, Hoda Fakhari, Ayse Lokmanoglu, Jessica Hullman, Steven Franconeri, Nicholas Diakopoulos, Erik C. Nisbet, and Matthew Kay. Swaying the public? Impacts of election forecast visualizations on emotion, trust, and intention in the 2022 U.S. midterms. In *Proceedings of the IEEE Visualization and Visual Analytics Conference*, 2023. DOI: 10.1109/TVCG.2023.3327356. **Best Paper Award (top 1%)**.

[3] **Fumeng Yang**, Chloe Mortenson, Nicholas Diakopoulos, Erik C. Nisbet, and Matthew Kay. In dice we trust: Uncertainty displays for maintaining trust in election forecasts over time. 2023. Under review at CHI.

[4] **Fumeng Yang**, Maryam Hedayati, and Matthew Kay. Subjective probability correction for uncertainty representations. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, 2023. DOI: 10.1145/3544548.3580998. **Best Paper Honorable Mention Award (top 5%)**.

[5] **Fumeng Yang**, Zhuanyi Huang, Jean Scholtz, and Dustin L. Arendt. How do visual explanations foster end users' appropriate trust in machine learning? In *Proceedings of the ACM Conference on Intelligent User Interfaces*, 2020. DOI: 10.1145/3377325.3377480. **Best Paper Honorable Mention Award (top 1%)**.

[6] **Fumeng Yang***, Caitlyn M. McColeman*, Timothy F. Brady, and Steven Franconeri. Rethinking the ranks of visual channels. In *Proceedings of the IEEE Visualization and Visual Analytics Conference*, 2021. DOI: 10.1109/TVCG.2021.3114684. **Best Paper Honorable Mention Award (top 5%)**. *equal contribution.

[7] 2020 Election Forecast. https://projects.fivethirtyeight.com/2020-election-forecast.

[8] Lane Harrison, **Fumeng Yang**, Steven Franconeri, and Remco Chang. Ranking visualizations of correlation using Weber's law. In *Proceedings of the IEEE Visualization and Visual Analytics Conference*, 2014. DOI: 10.1109/TVCG.2014.2346979.

[9] **Fumeng Yang**, Lane Harrison, Ronald A. Rensink, Steven Franconeri, and Remco Chang. Correlation judgment and visualization features: A comparative study. *IEEE Transactions on Visualization and Computer Graphics*, 2018. DOI: 10.1109/TVCG.2018.2810918.

[10] Brian D. Ondov, **Fumeng Yang**, Matthew Kay, Niklas Elmqvist, and Steven Franconeri. Revealing perceptual proxies with adversarial examples. In *Proceedings of the IEEE Visualization and Visual Analytics Conference*, 2020. DOI: 10.1109/TVCG.2020.3030429.

[11] **Fumeng Yang**, Yuxin Ma, Lane Harrison, James Tompkin, and David H. Laidlaw. How can deep neural networks aid visualization perception research? In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, 2023. DOI: 10.1145/3544548.3581111.

[12] **Fumeng Yang**, Jing Qian, Johannes Novotny, David Badre, Cullen D. Jackson, and David H. Laidlaw. A virtual reality memory palace variant aids knowledge retrieval from scholarly articles. *IEEE Transactions on Visualization and Computer Graphics*, 2020. DOI: 10.1109/TVCG.2020.3009003.

[13] **Fumeng Yang**, James Tompkin, Lane Harrison, and David H. Laidlaw. Visual cue effects on a classification accuracy estimation task in immersive scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, 2022. DOI: 10.1109/TVCG.2022.3192364.

[14] Jing Qian, Arielle Chapin, Alexandra Papoutsaki, **Fumeng Yang**, Klaas Nelissen, and Jeff Huang. Remotion: A motion-based capture and replay platform of mobile device interaction for remote usability testing. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2018. DOI: 10.1145/3214280.

[15] Jing Qian, Meredith Young-Ng, Xiangyu Li, Angel Cheung, **Fumeng Yang**, and Jeff Huang. Portalware: A smartphone-wearable dual-display system for expanding the free-hand interaction region in augmented reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2020. DOI: 10.1145/3334480.3383079.

[16] Yuan Cui, Lily W. Ge, Yiren Ding, **Fumeng Yang**, Lane Harrison, and Matthew Kay. Adaptive assessment of visualization literacy. In *Proceedings of the IEEE Visualization and Visual Analytics Conference*, 2023. DOI: 10.1109/TVCG.2023.3327165.