Last Updated: May 2024

Yunhao (Jerry) Zhang

University of California, Berkeley, Haas School of Business

2220 Piedmont Ave zyhjerry@mit.edu

Berkeley, CA 94720 https://www.yunhaojerryzhang.com/

EDUCATION

2017-2023	Massachusetts Institute of Technology
	MIT Sloan School of Management
	Ph.D. in Management
2012-2016	University of California, Berkeley
	B.A. in Economics and B.A. in Statistics

PROFESSIONAL

2023-2025 (Expected)	Postdoctoral Fellow
	Psychology of Technology Institute (co-sponsored by Berkeley Haas and USC Marshall)
	Advisors: Juliana Schroeder and Nate Fast

RESEARCH INTERESTS

Wisdom of Crowds, Collective Intelligence, Mechanism underlying Belief-updating, Consumer Perceptions of AI, Misinformation, Judgment & Decision Making

PUBLICATIONS

1. **Zhang, Y.**, & Gosline, R. R. (2023). People's Perceptions (and Bias) Toward Creative Content Generated by AI (ChatGPT-4), Human Experts, and Human-AI Collaboration. *Judgment and Decision Making*, Volume 18, 2023 , e41.
   DOI: https://doi.org/10.1017/jdm.2023.37

2. **Zhang, Y.**, & Rand, D. G. (2023). Sincere or motivated? Partisan bias in advice-taking. *Judgment and Decision Making*, Volume 18 , 2023 , e29
   DOI: https://doi.org/10.1017/jdm.2023.28

3. Arechar, A. A., Allen, J., Berinsky, A. J., Cole, R., Epstein, Z., Garimella, K., **Zhang Y.**, ... & Rand, D. G. (2023). Understanding and combatting misinformation across 16 countries on six continents. *Nature human behaviour*, 1-12.

4. Pennycook, G., McPhetres, J., **Zhang, Y.**, Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological science*, 31(7), 770-780.

5. Holtz, D., Zhao, M., Benzell, S. G., Cao, C. Y., Rahimian, M. A., Yang, J., **Zhang Y**, ... & Aral, S. (2020). Interdependence and the cost of uncoordinated responses to COVID-19. *Proceedings of the National Academy of Sciences*, 117(33), 19837-19843.

6. Cao, C., Cao, X., Cashman, M., Kumar, M., Timoshenko, A., Yang, J., **Zhang Y**,. ... & Wernerfelt, B. (2019). How do successful scholars get their best research ideas? An exploration. *Marketing Letters*, 30, 221-232.

WORKING PAPERS

1. Leveraging Advice-taking and Kernel Density Estimation to Identify A Cluster of Experts and Improve Wisdom of Crowds
   (under review at *Management Science*)

2. Self-Persuasion Does Not Imply Self-Deception
   (with David G. Rand, under review at *Cognition*)

3. Understanding the Dynamics of Aversion and Appreciation for Artificial Intelligence
   (with Renée Gosline, under review at *OBHDP*)
   (previously titled *Understanding Algorithm Aversion: When Do People Abandon AI After Seeing It Err?*)

4. What Predicts Accuracy?
   (with Don A. Moore, draft under preparation)

5. Predicting Others' Predictions: Leveraging Meta-Prediction Accuracy to Improve Social Influence in Binary Classification Problems
   (with Eaman Jahani, Douglas Guilbeault, Juliana Schroeder, draft under preparation)

6. The Revealed Confidence Algorithm: Leveraging Advice-taking to Identify Experts and Improve Wisdom of Crowds
(Working Paper – No intention to submit immediately because I have two better papers building on the concepts introduced in this paper.)

7. A Boundedly Rational Model of the Distance Effect in Advice-taking (Draft available upon request)

CONFERENCE AND SEMINAR PRESENTATIONS

Leveraging Advice-taking and Kernel Density Estimation to Identify A Cluster of Experts and Improve Wisdom of Crowds

- ACM Collective Intelligence Conference 2024, Boston, Jun 2024
- Informs SJDM Annual Meeting , November 2023

The Revealed Confidence Algorithm: Leveraging Advice-taking to Identify Experts and Improve Wisdom of Crowds

- Informs Advances in Decision Analysis Conference, June 2022
- MIT Human Cooperation Lab, Sep 2021
- ACM Collective Intelligence Conference 2021, Virtual, Jun 2021
- Max Planck Institute Center for Adaptive Rationality, Virtual, May 2021
- Informs SJDM Annual Meeting 2020, Virtual, Dec 2020
- MIT Conference on Digital Experimentation 2020, Virtual, Nov 2020
- MIT Behavioral Research Lab, Virtual, Oct 2020
- MIT Behavioral Economics Lunch, Virtual, Oct 2020
- Informs Marketing Science Conference 2020, Virtual, Jun 2020
- MIT Human Cooperation Lab, Virtual, Jun 2020
- MIT Sloan Marketing Seminar, Cambridge, MA, Oct 2019

Sincere or Motivated? Partisan Bias in Advice-taking

- University of Pennsylvania NoBeC (Norms and Behavioral Change) Talks for Early Career Researchers, Virtual, May 2022

- Informs Marketing Science Conference 2021, Virtual, Jun 2021
- UCL Affective Brain Lab, Virtual, April 2021
- MIT Behavioral Research Lab, Virtual, April 2021
- MIT Sloan Marketing Seminar, Virtual, Mar 2020
- MIT Human Cooperation Lab, Virtual, Feb 2021

Understanding Algorithm Aversion: When Do People Abandon AI After Seeing It Err?

- Informs Marketing Science Conference 2022, Virtual, June 2022
- MIT Human Cooperation Lab, Virtual, June 2022
- MIT Behavioral Research Lab, March 2022

GRANTS AND AWARDS

- AMA-Sheth Doctoral Consortium Fellow, 2022
- ADA Best PhD Incubator talk finalists, 2022
- ISMS Doctoral Consortium Fellow, 2020
- MIT Sloan School of Management Fellowship, 2017-2022

TEACHING EXPERIENCE

MBA Courses at MIT Sloan School of Management

- Applied Behavioral Economics (TA, Spring 2021, Spring 2022)
- Consumer Behavior (TA, Spring 2020, Spring 2022)
- Branding (TA, Spring 2021)

REFERENCES

Drazen Prelec (co-chair)
Professor of Management
Professor of Brain and Cognitive Sciences
Professor of Economics

David G. Rand (co-chair)
Professor of Management Science
Professor of Brain and Cognitive Sciences
drand@mit.edu

dprelec@mit.edu

ABSTRACTS FOR WORKING PAPERS

1. ***The Revealed Expertise Algorithm: Leveraging Advice-taking to Identify Experts and Improve Wisdom of Crowds***

   Identifying the experts within a crowd may help further improve the wisdom of crowds. I propose a new Revealed Confidence (RC) algorithm that uses the "RC measure", which is a scaled amount of belief updating given numerical advice (i.e., the group mean), as a proxy for prior variance to better reflect the relative expertise of each agent in a crowd. The intuition, which we confirm both theoretically and empirically, is that those who are less swayed by the group mean tend to be more accurate in their initial judgment. Therefore, using inverse-variance weighting with the RC measures as the variance inputs outperforms the existing wisdom-of-crowds methods by over-weighting the more accurate initial judgments in the aggregation. Crucially, we demonstrate that while self-reported confidence reflects *one's feeling of uncertainty given one's available information*, advice-taking reveals *the amount of information* one has and has not taken into account in their initial judgment. Therefore, the RC algorithm is able to successfully identify the experts, even when self-reported confidence fails. In addition, I show that the RC algorithm improves Wisdom of Crowds in a context where people might be biased (e.g., right-leaning Republicans and left-leaning Democrats answer political trivia questions).

2. ***Understanding Algorithm Aversion: When Do People Abandon AI After Seeing It Err?***

   Technological advancements have provided consumers the option to choose services provided by "artificial Intelligence" (AI) or human agents. Previous research (Dietvorst et al. 2015) has shown that consumers display "algorithm aversion" after seeing AI err. We further explore this phenomenon by testing the boundaries of this effect. In a series of experiments examining preference for an AI or a human forecaster in statistical prediction tasks, we find that (1) when participants are informed of either the AI's or the human forecaster's previous error, algorithm aversion disappears as the AI's (human forecaster's) error in the feedback becomes smaller (larger); (2) when the feedback suggests both the human forecaster and the AI have the same previous error, participants do not abandon the AI. These results are a consequence of participants quite rationally updating the relative competence of the AI and the human forecaster based on the accuracy cue. Furthermore, we find that algorithm aversion does not happen if an accuracy cue suggests the AI performs worse than one's expectation, but rather when the feedback indicates the AI's error is larger than one's expected human forecaster's error. This suggests that the perceived relative competence of the AI and the human forecaster ultimately determines participants' preference. Overall, our findings suggest that people tolerate AI's imperfection to a degree greater than previously thought, but firms should strive to improve the accuracy of their AI as this affects algorithm acceptance.

3. *A Boundedly Rational Model of the Distance Effect in Advice-taking*

Since the assumption that agents can follow perfect Bayesian computation is often impractical, I develop a semi-Bayesian belief-updating model to characterize the relationship among stated confidence, uncertainty, expertise, and advice-taking. The model shows that the amount of advice-taking (weight on advice) reveals both first-order uncertainty (e.g., the width of the stated confidence interval) and second-order uncertainty (e.g., uncertainty about the width of the stated confidence interval). The model is able to reconcile two important empirical phenomena. First, we demonstrate that even though agents can state a high confidence (i.e., low first-order uncertainty), they may put a large weight on the advice in belief-updating if their estimate of their stated confidence is imprecise (i.e., large second-order uncertainty due to their lack of information). Second, we show that the distance effect (i.e., the weight on advice tends to decrease as the distance between the initial estimate and the advice increases), a widely documented empirical pattern in advice-taking, can be a consequence of people updating their beliefs following a semi-Bayesian updating heuristics given their cognitive limitation. We discuss the implication of these findings in the context of the Revealed Expertise (RC) Algorithm. For example, since a particular realization of an outcome can be random, experts does not always have a lower first-order uncertainty (e.g., not sure whether a coin flip lands on heads or tails), but they tend to have low second-order uncertainty (e.g., very certain that a fair coin has 50\% chance of landing on heads). Therefore, the RC algorithm is able to correctly identify experts even when self-reported confidence fails. Furthermore, the extent the advice potentially contains "surprising" information is a function of the distance between an initial answer and the advice. Therefore, distant advice such as the group mean serves as a great benchmark to reveal the relative expertise among agents when applying the RC algorithm.