Last Updated: September 2022

Yunhao (Jerry) Zhang

MIT Sloan School of Management

100 Main St, E62-378 zyhjerry@mit.edu

Cambridge, MA 02142 https://www.yunhaojerryzhang.com/

EDUCATION

2017-2023 (Expected)	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

RESEARCH INTERESTS

Wisdom of Crowds, Confidence & Uncertainty, Advice-taking, (Identity-based) Motivated Reasoning & Inter-group Collaboration, Algorithm Aversion & Acceptance, Misinformation, Judgment & Decision Making, Behavioral & Experimental Economics.

PUBLICATIONS

1. Pennycook, Gordon, Jonathon McPhetres, **Yunhao Zhang**, Jackson G. Lu, and David G. Rand. "Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention." ***Psychological science*** 31, no. 7 (2020): 770-780.

2. Holtz, David, Michael Zhao, Seth G. Benzell, Cathy Y. Cao, Mohammad Amin Rahimian, Jeremy Yang, Jennifer Allen et al. "Interdependence and the cost of uncoordinated responses to COVID-19." ***Proceedings of the National Academy of Sciences*** 117, no. 33 (2020): 19837-19843.

3. Cao, Cathy, Xinyu Cao, Matthew Cashman, Madhav Kumar, Artem Timoshenko, Jeremy Yang, Shuyi Yu, **Jerry Zhang**, Yuting Zhu, and Birger Wernerfelt. "How do successful scholars get their best research ideas? An exploration." ***Marketing Letters*** 30, no. 3 (2019): 221-232.

WORKING PAPERS

1. The Revealed Expertise Algorithm: Leveraging Advice-taking to Identify Experts and Improve Wisdom of Crowds
   (Job Market Paper. Resubmit and under review at ***Management Science***)

2. Sincere or Motivated? Partisan Bias in Advice-taking (with David G. Rand) (Submitted to ***Nature Human Behavior***)

3. Understanding Algorithm Aversion: When Do People Abandon AI After Seeing It Err? (with Renee Gosline) (Draft available upon request. Preparing for Submission to ***Management Science***)

4. A Boundedly Rational Model of the Distance Effect in Advice-taking (Draft available upon request. Preparing for Submission to ***Management Science***)

5. Understanding and reducing online misinformation across 16 countries on six continents (with Antonio Alonso Arechar, Jennifer Nancy Lee Allen, Rocky Cole, Ziv Epstein, Kiran Garimella, Andrew Gully, Jackson G Lu, Robert M Ross, Michael Stagnaro, Gordon Pennycook, David Rand) (R&R at ***Nature Human Behavior***)

CONFERENCE AND SEMINAR PRESENTATIONS

The Revealed Expertise 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 (with David G. Rand)

- 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? (with Renee Gosline)

- 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
dprelec@mit.edu

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

Renee R. Gosline
MIT Sloan Senior Lecturer
MIT IDE Principle Research Scientist
rgosline@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 Expertise (RE) algorithm that uses the "RE 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 RE 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 RE algorithm is able to successfully identify the experts,

even when self-reported confidence fails. In addition, I show that the RE 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). Nevertheless, I still propose a pre-registered method in which we measure subjects' bias in advice-taking to calibrate the RE measures and further improve the algorithm's performance.

2. *Sincere or Motivated? Partisan Bias in Advice-taking*

Political divisions have become a central feature of modern life. Here, we ask whether these divisions affect advice-taking from co- and counter-partisans. In an incentivized task assessing the accuracy of non-political news headlines, we find partisan bias in advice-taking: Democratic participants are less swayed by (accurate) information that comes from Republicans compared to the same information from Democrats (Republican participants display no such bias). We then adjudicate between two possible mechanisms for this biased advice-taking: a preference-based account, where participants are motivated to take less advice from counter-partisans because doing so is unpleasant; versus a belief-based account, where participants sincerely believe co-partisans are more competent at the task (even though this belief is incorrect). To do so, we examine the impact of a 1000-fold increase in the stakes, which should increase accuracy motivations (and thereby reduce the relative impact of partisan motivations). We find that increasing the stakes does not reduce biased advice-taking, hence no evidence to support the bias is driven by preference. Consistent with the belief-based account, we find that Democratic participants (incorrectly) believe their co-partisans are better at the task, and this incorrect belief is much less severe among Republican participants. Further supporting the notion that the stated beliefs are sincere, raising the stakes of the belief elicitation of relative partisan competence does not affect the stated beliefs. Finally, participants – instead of ignoring the feedback – update in favor of their counter-partisans given feedback that suggests counter-partisans are competent. The implication of our study is that political divisions perhaps are not driven by a preference-based distaste of counter-partisans but by people having incorrect beliefs about counter-partisans’ competence (e.g., due to a lack of access to news that sustains a positive image of counter-partisans).

3. *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.

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

(Note: this computational model was originally written for the paper about the Revealed Expertise Algorithm. Following the suggestions by the referees at Management Science, this model is now a separate paper under preparation for submission.)

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 (RE) 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 RE 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 RE algorithm.

5. ***Understanding and reducing online misinformation across 16 countries on six continents***

The spread of misinformation online is a global problem that requires global solutions. To that end, we conducted an experiment in 16 countries across 6 continents (N = 33,480) to investigate predictors of susceptibility to misinformation and interventions to combat misinformation. In every country, participants with a more analytic cognitive style and stronger accuracy-related motivations were better at discerning truth from falsehood; valuing democracy was also associated with greater truth discernment whereas political conservatism was negatively associated with truth discernment in most countries. Subtly prompting people to think about accuracy was broadly effective at improving the veracity of news that people were willing to share, as were minimal digital literacy tips. Finally, crowdsourced accuracy evaluation was able to differentiate true from false headlines with high accuracy in all countries. The consistent patterns we observe suggest that the psychological factors underlying the misinformation challenge are similar across the globe, and that similar solutions may be broadly effective.