

PSYC GU4286: The Accuracy of Human Judgment Seminar Syllabus

Columbia University, New York

Fall 2023

09/5/2023 – 12/22/2023

Instructor: Dr. John Wilcox

Email: wilcoxje@stanford.edu (provisional email address)

Day/Time: --

Classroom: --

Office hour location: --

Office hour time: Arranged by prior appointment

Anonymous feedback form: <https://forms.gle/eTDEf25KVsoze6Wc6>

Course Overview:

Course Description: We all make judgments about what is true or false, probable or improbable. Additionally, we all use these judgments to inform important decisions: who to marry, what career to pursue, where to live, what medications to take, which theories to accept and who to trust, to take a few of countless many examples. What differentiates us, however, is how accurate these judgments are: research has shown that some individuals and groups are much more accurate than others, and you might be surprised at which variables do (or do not) correlate with this accuracy.

In this course, we will investigate how to understand, measure and improve the accuracy of human judgment. Topics covered include the following: the concept and measurement of judgmental accuracy, studies assessing judgmental accuracy across various domains, the accuracy of our metacognition (that is, of our thoughts about our thoughts), heuristics as potential causes of inaccuracy, evolutionary explanations of the inaccuracy of human judgment and studies revealing means by which to potentially improve the accuracy of our judgments. We will also explore applications to real-world contexts, including law, medicine, geopolitics and the concept of “expertise”.

Course Aims: The course’s aims are threefold: 1) to acquire introductory-level knowledge about the aforementioned topics regarding human judgment, 2) to develop generalizable scientific thinking skills and 3) to develop written and oral communication skills. Each of these aims in turn gives rise to the following sub-aims.

- 1) The acquisition of topical knowledge:
 - To gain familiarity with some prominent measures of judgmental accuracy
 - To acquire familiarity with evidence indicating accuracy or inaccuracy across various domains, including law, medicine and politics, as well as metacognition in general
 - To better understand the role of heuristics in human thinking

- To gain familiarity with competing evolutionary explanations of human inaccuracy
 - To gain familiarity with some variables that correlate with, or improve, the accuracy of human judgment
- 2) The development of scientific thinking skills:
- To further develop the ability to assess studies along specific dimensions of scientific rigor, particularly the following:
 - *measurement validity*: whether a study's measure(s) can be interpreted as accurate measures of the intended constructs
 - *internal validity*: whether effects can be attributed to their putative causes
 - *external validity*: whether effects can be generalized to the intended contexts of interest
 - To further develop the ability to generate study designs that are appropriate for investigating important research questions
- 3) The development of written and oral communication skills:
- To further develop oral communication skills through giving a presentation
 - To further develop written communication skills with short written assignments and a longer report

Importantly, because the course focuses on developing rigorous thinking skills, do not be surprised if you receive more critical feedback than you might encounter in particular other courses. And please: do not take it personally or be discouraged. Everyone receives similarly critical feedback, it may not necessarily severely impact your grade, and the purpose is to help you think in ways that will serve you well for this course and—hopefully—for your life more broadly, as students from other universities can tell you.

Additionally, this course welcomes diversity—not only in terms of gender, race, sexual orientation, religion, ability and so on—but also in terms of viewpoints. This sometimes means tolerating and ideally respecting those whose views we disagree with—and sometimes disagree with strongly. If you feel too uncomfortable discussing arguments that, for example, racism explains higher conviction rates among particular minorities, then you should perhaps consider not taking this course. Conversely, if you feel too uncomfortable discussing arguments that racism *does not* explain higher conviction rates among particular minorities, then you also should perhaps consider not taking this course. You do not have to agree with everything you read, but open-mindedness, humility and respectful disagreement are necessary in this course.

Lastly, and relatedly, you are welcome to disagree with the instructor (Professor John Wilcox). Sometimes students have written excellent essays arguing against his views and have received top grades; conversely, sometimes students have written poor essays supporting his views and have received low grades. What matters most is clarity and scientific rigor, but this requires also sufficient attentiveness to arguments against your perspectives too.

Course Topics and Schedule:

Structure of the Sessions: Each session will focus on a topic to do with the accuracy of human judgment. For what follows, a brief description of each session is provided, along with focus questions and a set of

readings. The focus questions are there partly because not everything you read will be important or remembered, and so the questions can guide your attention, reading and reflections to issues that are especially important or will be discussed in the sessions. Students, especially graduates, may wish to ignore these questions if they find that the questions are unhelpful or unduly restrict their attention. The readings for the first week are at times philosophical and hypothetical, but do not be fooled: we will be diving into empirical scientific work, starting in the second week with our exploration of judgmental accuracy across cultures.

September 8th: Session 1 – The accuracy of human judgment: its importance, conceptualization and measurement

Since the topic of the course is the accuracy of human judgment, we will start the course by reflecting on what judgmental accuracy means and why it is important.

Focus questions:

- 1) Why might one think judgmental accuracy is an important topic?
- 2) What are some of the different measures of judgmental accuracy?
- 3) What are the strengths and limitations of each measure?

Readings:

- 1-2. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Sections 1.1 and 1.2, and Chapter 2.
- 3-4. Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: a flaw in human judgment*. Hachette UK. Introduction and Chapter 5.

September 15th: Session 2 – Evidence of inaccuracy: Judgmental accuracy across cultures

A range of studies have reported that judgmental accuracy varies across cultures. Not only is this interesting for assessing the judgmental accuracy of everyday persons in different societies, but it is also interesting to consider what explains any putative cultural differences.

Focus questions:

- 1) Why might one think judgmental accuracy across cultures is an important topic?
- 2) What is the evidence bearing on how accurate human judgment is across cultures?
- 3) How good is that evidence, and how could better evidence be obtained?

Readings:

1. Yates, J. F. (2010). Culture and probability judgment. *Social and Personality Psychology Compass*, 4(3), 174-188.
2. Lechuga, J., & Wiebe, J. S. (2011). Culture and probability judgment accuracy: The influence of holistic reasoning. *Journal of cross-cultural psychology*, 42(6), 1054-1065.

3. Yates, J. F., Lee, J. W., Sieck, W. R., Choi, I., & Price, P. C. (2002). Probability judgment across cultures. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 271–291). Cambridge University Press.

September 22nd: Session 3 – Evidence of inaccuracy: Judgmental accuracy in law

Last week, we explored the accuracy of everyday individuals across cultures. This week, we will explore the accuracy of everyday individuals in a special context: law. The law and the institutions that administer it would ideally play an important part in maintaining a just and flourishing society. But how accurate are the judgments of individuals—and specifically juries—that can sentence some individuals to death while pronouncing others innocent?

Focus questions:

- 1) Why might one think judgmental accuracy in law is an important topic?
- 2) What is the evidence bearing on how accurate human judgment is in law?
- 3) How good is that evidence, and how could better evidence be obtained?

Readings:

1. Gross, S. R., O'Brien, B., Hu, C., & Kennedy, E. H. (2014). Rate of false conviction of criminal defendants who are sentenced to death. *Proceedings of the National Academy of Sciences*, *111*(20), 7230–7235. <https://doi.org/10.1073/pnas.1306417111>
2. Gross, S. R. (2017). What We Think, What We Know and What We Think We Know about False Convictions. *Ohio State Journal of Criminal Law*, *14*(2), 753–786.
3. Cassell, P. G. (2018). Overstating America's Wrongful Conviction Rate: Reassessing the Conventional Wisdom about the Prevalence of Wrongful Convictions. *Arizona Law Review*, *60*(4), 815–864.

September 29th: Session 4 – Evidence of inaccuracy: Judgmental accuracy of experts vs. non-experts (Part I: epidemiology and the COVID-19 pandemic)

So far, we have explored the accuracy of primarily everyday individuals. However, society also has various kinds of “experts” who are widely supposed to have more accurate opinions about their domain of expertise. Over the next two weeks, we will investigate evidence comparing the accuracy of judgments among experts and non-experts in particular domains, focusing particularly on epidemiology and then on politics and social science. We will use the COVID-19 pandemic as a case study this week, examining specifically a televised exchange between a non-expert who argued for an immediate lockdown in the UK and an expert who argued to the contrary. In retrospect, the expert claimed that the delayed lockdown was a mistake, but this was only 3 months later when approximately 39,048 individuals had died from the COVID-19 virus. Such instances warrant reflection about how expertise relates to judgmental accuracy and trustworthiness more generally, both in theory and in practice.

Focus questions:

- 1) Why might one think the judgmental accuracy of experts is an important topic?
- 2) What is the evidence bearing on how accurate expert and non-expert judgment is in epidemiology?
- 3) How good is that evidence, and how could better evidence be obtained?

Readings:

1. Recchia, G., Freeman, A. L., & Spiegelhalter, D. (2021). How well did experts and laypeople forecast the size of the COVID-19 pandemic? *PloS One*, *16*(5), e0250935.
2. dan. (2020, June 2). A Preliminary Look at Metaculus and Expert Forecasts. <https://www.metaculus.com/news/2020/06/02/LRT/>
3. Coronavirus Special: Are We Doing Enough? – Channel 4 News. (2020). *YouTube*. <https://youtu.be/C98FmoZVbjs?t=590> (9:50-23:34)
4. Coronavirus: Lockdown Delay “Cost a Lot of Lives”, Says Science Adviser—BBC News. (2020, June 7). *BBC News*. <https://www.bbc.com/news/uk-politics-52955034>

October 6th: Session 5 – Evidence of inaccuracy: Judgmental accuracy of experts vs. non-experts (Part II: politics and social science)

This week, we will explore the accuracy of experts in social science and in politics. Of course, a related line of research has found that some individuals are remarkably accurate in geopolitical forecasting, yet we will examine this research in some subsequent weeks.

Focus questions:

- 1) Why might one think judgmental accuracy in social science or politics is an important topic?
- 2) What is the evidence bearing on how accurate human judgment is in social science and politics?
- 3) How good is that evidence, and how could better evidence be obtained?

Readings:

1. The Forecasting Collaborative, "Insights into the accuracy of social scientists' forecasts of societal change." *Nature human behaviour* 7, no. 4 (2023): 484-501.
- 2-3. Tetlock, P. (2005). *Expert political judgment: How good is it? How can we know?* Princeton University Press. Chapters 1 and 2.

October 13th: Session 6 – Evidence of inaccuracy (and accuracy!): Judgmental accuracy in intelligence and government decision-making

This week, we move on from comparisons of experts and non-experts to instead consider the accuracy of particular groups of individuals, starting with intelligence analysts and government

officials. Intelligence analysts are frequently the eyes through which government officials see the world and make decisions in response to it. Such analysts and officials provide judgments about important topics, such as the origins of COVID-19, the threats posed by other nations and the prospects of a potential military strategy's success. How accurate, then, are the judgments which intelligence analysts and government officials produce?

Focus questions:

- 1) Why might one think judgmental accuracy in government decision-making is an important topic?
- 2) What is the evidence bearing on how accurate human judgment is in government decision-making?
- 3) How good is that evidence, and how could better evidence be obtained?

Readings:

1. Mandel, D. R., & Irwin, D. (2021). Tracking accuracy of strategic intelligence forecasts: Findings from a long-term Canadian study. *Futures & Foresight Science*, 3(3-4), 1-13.
2. Friedman, J. A. (working paper). "Results from a Large-Scale Study of Overconfidence among National Security Officials".
3. Mandel, D. R., & Barnes, A. (2014). Accuracy of forecasts in strategic intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, 111(30), 10984–10989. <https://doi.org/10.1073/pnas.1406138111>.

October 20th: Session 7 – Evidence of inaccuracy: Judgmental accuracy in medicine

Health is essential for the functioning of society: we literally cannot live without it. Medical professionals then have an important duty in maintaining the health of society. Consequently, this week, we will examine the accuracy of judgments from medical professionals.

Focus questions:

- 1) Why might one think judgmental accuracy in medicine is an important topic?
- 2) What is the evidence bearing on how accurate human judgment is in medicine?
- 3) How good is that evidence, and how could better evidence be obtained?

Readings:

1. Graber, M. L. (2013). The incidence of diagnostic error in medicine. *BMJ quality & safety*, 22(Suppl 2), ii21-ii27.
2. Lam, J. A., & Edward Feller, M. D. (2020). Are We Right When We're Certain? Overconfidence in Medicine. *Rhode Island Medical Journal*, 103(2), 11-12.
3. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Section 3.3.

October 27th: Session 8 – Evidence and causes of inaccuracy: Metacognitive inaccuracy

We have seen a range of studies, many of which indicate that humans are inaccurate in their judgments—and with harmful consequences. We will now turn to examine some potential causes (or enablers) of this inaccuracy, starting with inaccurate metacognition—that is, in this context, inaccurate judgments about the accuracy of our judgments. This week, we will review some of the evidence for metacognitive inaccuracy, as well as some explanations of it.

Focus questions:

- 1) What is metacognitive inaccuracy?
- 2) What is the evidence for metacognitive inaccuracy, and how good is it?
- 3) What are some potential explanations of metacognitive inaccuracy?

Readings:

1. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Chapter 4.
2. Händel, M., & Dresel, M. (2018). Confidence in performance judgment accuracy: The unskilled and unaware effect revisited. *Metacognition and Learning*, 13, 265-285.
3. Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), 1121-1134.

November 3rd: Session 9 – Causes of inaccuracy: Heuristics and biases

Inaccurate metacognition can partially explain why judgmental inaccuracy persists, but it does not necessarily explain how it arises in the first place. To explain the occurrence of judgmental inaccuracy, we might instead turn to a well-studied subject in the judgment and decision-making literature: heuristics and biases, the topic of this week.

Focus questions:

- 1) What is a bias, and what is a heuristic? How are the two different?
- 2) What are some of the main heuristics?
- 3) What are different perspectives on heuristics advanced by figures such as Kahneman compared to figures such as Gigerenzer?

Readings:

1. Baron, J. (2014). Heuristics and biases. *The Oxford handbook of behavioral economics and the law*, 3-27.
2. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Section 5.4.
3. Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in cognitive science*, 1(1), 107-143.

November 10th: Session 10 – Causes of inaccuracy: Evolutionary theories of inaccuracy

So we have seen evidence about how accurate human judgment is, as well as heuristics that sometimes explain its inaccuracy. This week, we will take a step even further back and consider evolutionary explanations for our putative cognitive defects at a higher level. To do that, we will take as a starting point an influential account of human reason, called the “interactionist approach”, developed by cognitive scientists Hugo Mercier and Dan Sperber.

Focus questions:

- 1) What are the two main evolutionary accounts of reason that are discussed by Mercier and Sperber?
- 2) What are the strengths and limitations of those accounts?

Readings:

1. Mercier, H., & Sperber, D. (2019). Précis of the enigma of reason. *Teorema: Revista Internacional de Filosofía*, 38(1), 69-76.
2. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Chapter 6.
3. Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press, Introduction.

November 17th: Session 11 – Presentations

This week, students will give presentations on the topic of their research papers, as per the assessment structure detailed below.

November 24th: No Classes – University Holiday

Go party.

December 1st: Session 12 – Correlates of accuracy: Overviews and early research

So far, this course may have painted a grim picture of our psychologies: we have looked at a lot of evidence that human judgment is often (though not always) inaccurate, that we often do not realize it, that our accuracy is often systematically compromised by heuristics and that our evolutionary history may have permitted these epistemic defects for one reason or another. This week, however, we take a more positive turn: a lot of research has shown both that human judgment can be impressively accurate and that specific variables correlate with, or improve, our accuracy. The rest of the course will be devoted to examining this more positive message from the science of human judgment.

Focus questions:

- 1) What are some of the main correlates of judgmental accuracy?
- 2) How convincing is the evidence about what improves judgmental accuracy in general?
- 3) How can we get better evidence?

Readings:

1. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Chapter 7.
2. Mellers, B., Stone, E., Atanasov, P., Rohrbaugh, N., Metz, S. E., Ungar, L., Bishop, M. M., Horowitz, M., Merkle, E., & Tetlock, P. (2015). The psychology of intelligence analysis: Drivers of prediction accuracy in world politics. *Journal of experimental psychology: applied*, 21(1), 1.
3. Tetlock, P. E., & Gardner, D. (2016). *Superforecasting: The art and science of prediction*. Random House, Chapter 1.

December 8th: Session 13 – Correlates of accuracy: Recent research

This week, we will continue our investigation of research into the variables that can improve or correlate with accuracy.

Focus questions:

- 1) What are some of the main correlates of judgmental accuracy covered in this week's readings?
- 2) How convincing is this evidence about what improves judgmental accuracy in general?
- 3) How can we get better evidence?

Reading:

1. Karvetski, C. W., Meinel, C., Maxwell, D. T., Lu, Y., Mellers, B. A., & Tetlock, P. E. (2022). What do forecasting rationales reveal about thinking patterns of top geopolitical forecasters?. *International Journal of Forecasting*, 38(2), 688-704.
2. Chang, W., Atanasov, P., Patil, S., Mellers, B. A., & Tetlock, P. E. (2017). Accountability and adaptive performance under uncertainty: A long-term view. *Judgment and Decision Making*, 12(6), 610-626.
3. Wilcox, J. (2022). Human Judgment: How Accurate is it, and How Can it Get Better? *Springer*. Chapter 8.

Assessment Structure and Schedule:

Students' overall course grades will be determined by the following assessments:

1. Each week:

In-class participation – 10%: Students must attend classes and they must contribute to each seminar. Practically, their contributions would ideally look like interested engagement with class discussions, and one (crude) operationalization of this is contributing at least one thoughtful comment or question in each class that one attends. That said, students can also miss one class for unspecified reasons (e.g. taking care of their mental health).

2. One week after the relevant session, and two before session 7 (October 20th) and two before session 11 (November 17th):

Four short (200-700 word) methodological critiques and suggestions - 40% (Pass/Fail): To foster knowledge acquisition, scientific thinking skills and written communication skills, students must submit four short methodological critiques. Two of these must be submitted before session 7 and the other two must be submitted before session 11. Each methodological critique is essentially your response to the following prompt:

- 1) *Identify some methodological limitation(s) in one of the readings for a given week, and*
- 2) *Constructively suggest a methodological improvement that could assist in answering some important or interesting research question that is the same as (or related to) the original reading*

No two critiques may concern the same week, and **each critique is due one week after the respective session/topic which it concerns**. These critiques will be pass/fail and are easy to pass in order that students will not feel undue anxiety about submitting them. However, they will also be assigned hypothetical letter grades that indicate what grade they *would have received* had they been graded with letters. Most importantly, these critiques will serve as a guide to expectations and how students will be evaluated for the final research paper. So they should be taken seriously as opportunities to practice for the final research paper; otherwise, students may perform poorly on the paper and compromise their final grade.

3. October 27th, 11:59pm:

Short (500-1,500 word) research paper proposal - 5% (Pass/Fail): Students are expected to submit a short research paper proposal outlining: 1) the research question that they are investigating in the final research paper, 2) its importance and 3) what method(s) they will use to investigate it. More details about the final paper (and hence research paper proposal) can be found below. Note that this is not a research *methods* proposal which comprehensively outlines methods for an empirical study. Rather it is a research paper

proposal which aims merely to describe what the content of your final paper will be, where the content of your paper may or may not be a research methods proposal.

4. November 17th:

Presentations - 10% (Letter grade): To foster knowledge acquisition, scientific thinking skills and oral communication skills, students must give a 10-minute presentation on their research project which they will write about in their research paper.

5. December 13th, 11:59pm:

Final research paper (2,500-3,500 words) – 35% (Letter grade): The final research paper is meant to showcase your achievement of the course objectives: the acquisition of knowledge, the development of scientific thinking skills and the development of (in this case, written) communication skills. As such, you will be expected to produce a research paper that demonstrates some knowledge of judgmental accuracy, that uses scientific research skills (such as data analysis) to investigate a question and that clearly communicates its content. The research paper could look like the following: a new study design which proposes to collect data to test one or more hypotheses (although you would not necessarily collect this data); an analysis of some existing dataset to test some hypotheses; the proposal of a new theory, as well as some discussion of scientific methods to test it; or perhaps something else.

Policies and Accommodations:

Electronics Policy: The use of electronic devices is softly discouraged, but it is nevertheless permitted in class only for the sake of accessing readings or notes offline. If you need to use a device, please be respectful of your classmates and your instructor by only accessing the document for class and not using the device to message or to access the internet.

Academic Dishonesty: Academic honesty is taken very seriously. Columbia students commit to the Honor Code as follows: “I affirm that I will not plagiarize, use unauthorized materials, or give or receive illegitimate help on assignments, papers, or examinations. I will also uphold equity and honesty in the evaluation of my work and the work of others. I do so to sustain a community built around this Code of Honor.” All suspected cases of dishonest behavior will be reported to Student Conduct and Community Standards (SCCS) and may result in severe consequences.

Disability Services: If you require additional assistance with assignments or exams, please check in with the Office of Disability Services. More information is available at:
<https://health.columbia.edu/content/disability-services>

COVID-19 policies: We will comply with all University-mandated COVID-19 policies.