

## **Wharton Undergraduate Elective, MGT 276, Spring, 2016**

### **Probability Judgments: Forecasting in Business and Politics**

**Prof. Philip E. Tetlock** (Steinberg-Dietrich 3203: [tetlock@wharton.upenn.edu](mailto:tetlock@wharton.upenn.edu))

**Teaching Assistance: Welton Chang** [welton.chang@gmail.com](mailto:welton.chang@gmail.com)

#### **Course Goals**

This course will introduce students to a wide range of approaches for assessing and improving judgment under uncertainty, with special emphasis on improving probability judgments of possible futures.

A world-class poker player once said he could spot the serious players among talented amateurs because the serious players knew the difference between 40-60 bets and 60-40 bets. In other words, serious players are well calibrated: they are skilled at distinguishing degrees of confidence. It is a good idea for Wharton students to get their calibration workouts in low-risk settings, such as classrooms, before the real world starts delivering feedback less gently.

This course is intended to give students opportunities both to fine-tune their own judgment skills as well as to master and then weave together insights from scientific research on judgment. Students will have repeated opportunities to: (a) learn how to translate interesting questions into resolvable questions—and to translate vague hunches about the right answers into probability metrics that can be scored for accuracy; (b) explore methods of making team forecasts that are more than the sum of their individual-team-member parts; (c) explore approaches to motivating people to put aside goals that often impair accuracy—and focus solely on improving foresight.

The ultimate teaching goal is to bridge the gap between “private” and “public” knowledge—a gap captured by the distinction between knowing how to ride a bicycle and knowing Newton’s laws of motion. In other classes, students learn about formal models for guiding business decisions. In this class, they will be free to draw on whatever eclectic mix of formal models and intuition they choose. The focus will be on their bottom-line accuracy in sizing up real world problems and their adroitness, as individuals and as teams, in learning from feedback. I want students to launch themselves on a lifelong process of questioning how much trust they should place in knowledge claims, their own as well as those of experts whose opinions they value—be they pundits,, colleagues, business leaders or indeed professors.

#### **Phil Tetlock Bio:**

I have been running forecasting tournaments for almost as long as he has been a professor (a long time). So this course is a natural extension of my life’s work. In 2005, I wrote "Expert political judgment: How good is it? How can we know?" which tracked the accuracy of leading political experts between 1985 and 2005—and in 2015, I wrote “Superforecasting: The art and science of prediction.” And I was the Principal Investigator for Team Good Judgment, which won a series of geopolitical/economic forecasting tournaments (2010-2015) sponsored by the U.S. Government. My appointment at Penn is interdisciplinary. I hold the Annenberg University

Professorship, a PIK position that bridges Wharton and the School of Arts and Sciences. My full vita is on my university websites.

**Grading (pass-fail not permitted):** Grading philosophy is grounded in the well established research finding that frequent short testing produces deeper and longer-term learning than infrequent equivalent-duration testing. Grades will be based on

(a) 10-minute quizzes at the start of each class from weeks 2 through 11 (to ensure everyone is keeping up—35% of grade—but we will drop the two worst quiz grades);

(b) class and forecasting tournament participation (35%),

(c) a major presentation on the final day of class (30%);

**Week 1 (January 19). Judging judgment and keeping score:** Most forecasts in high-stakes debates are verbal, not numerical. Famous pundits tell us that x “might” or “could” happen—or there is a “distinct possibility” or “fair chance” of x. When we ask people to translate this vague verbiage into probabilities, their answers are erratic, from as low as 0.01 to as high as 0.80.

This course shows how we can all do better than that.

We ask you to do three things in preparing for the first class, so we can get a good running start:

Read: Chapter 1 and 2 “Superforecasting.” Crown, 2015.

Plus Chapter 9 for guidance on creating your own “superforecasting” team.

Sign up as forecaster at [gjopen.com](http://gjopen.com) and make forecasts on at least five questions (give it careful thought because your accuracy scores will be part of class participation grade—a large part if you so choose). Pick questions that will be resolved within length of this class—and pick a few that will resolve in the next month so you can get early “accuracy feedback. **Send your GJopen username to TA Welton Chang so that you will receive credit for your forecasts.**

**In Class Team Exercise:** Each team will pick a favorite pundit on a topic of interest. Why that person? What credibility cues do you use in judging him or her? What do you *really know* about his or her track record? Why might pundits prefer vague-verbiage forecasts? How might this preference for opacity affect the quality of debate in our society? Each team will present on their pundit to the class.

**Team Homework Exercise:** Teams will formulate and nominate five questions that they find particularly interesting and want to track during the course of the class. Teams will submit initial forecasts and updates on the questions weekly (by 5pm on Friday) to the TA.

**Week 2. Knowing what you know and you don’t: The perils of false knowledge.** How predictable do you tacitly assume a problematic situation to be (where along the “clouds vs

clocks” continuum does it fall)? And how do people so often manage to convince themselves that they know things about the world, and about the forecasting prowess of famous figures, that are unknowable given current evidence.

**Readings:** Chapters 1, 2 and 3 of Rosenzweig, P. (2006). *The halo effect and eight other business delusions that deceive managers*. New York; Free Press.

NFL “Deflategate” Report: pp.121-131

Iran Nuclear Capability National Intelligence Estimate Key Judgments: pp. 1-9

**Team Exercise:** Students will be introduced to an informal version of Precision Questioning that teaches them how to “disagree without being disagreeable.” Students will also explain three forecasts you made on gjoin.com with other team members—and start constructively critiquing each others’ forecasts (learn to disagree without being disagreeable). Make the materials you prepare part of your team participation file which you will need to submit at the end of the term for grading.

**Week 3. Deserved success or blessed by chance? Distinguishing Luck from Skill.** This session explores the challenges of disentangling luck vs. skill.

We will also discuss how our assumptions about the possibility of improving our judgment can become self-fulfilling prophecies. Are your implicit beliefs about the roles of luck and skill are holding them back from cultivating career-enhancing skills?

**Team Exercises:** Teams will learn to calculate the accuracy of their forecasts and learn to do “pre-mortems” on their forecasts (ways in which their forecasts might die).

**Readings:** Mauboussin, M. (2012). *The success equation: Untangling success and luck in business, sports, and investing*. Harvard Business Review Press. Chapters 1-3.

St. Louis Federal Reserve Bank Report: Education and Wealth, pp. 4-7.

Eugene Fama, Kenneth French, “Luck versus skill in the cross- section of mutual fund returns”, *Journal of Finance*, 2010, pp. 1915-1935

**Week 4: Treating your beliefs like testable propositions, not sacred possessions: Common pitfalls in hypothesis testing.**

Students will be introduced to the concept of Bayesian belief updating and to research that identifies the conditions under which we should expect forecasters to fall prey to the errors of either under-reacting to new evidence (belief perseverance) or of over-reacting (excessive volatility).

**Readings:** Chapters 4 through 8 in *Superforecasters*  
Chapters 4-7 in Rosenzweig’s *Business delusions*

**Team Exercise:** Teams will work on a belief updating problem developed by the Good Judgment Lab. Teams will also identify examples of evidence in real-world forecasting that would cause them either to under-react (missing “subtly diagnostic” evidence) or to over-react (getting “sucker-punched” by pseudo-diagnostic” evidence).

**Week 5 Hindsight bias—seeing through the narrativist fallacy. Just because pundits can offer colorful explanations for something does not mean they could have predicted it.**

**Readings:**

Bazerman, M. Predictable surprises. Harvard Business Review.

Tetlock, P.E. (2005). Chapters 6 and 7 in Expert political judgment. Princeton University Press

**Team Exercise:** Students will try to reconstruct the expectations they had for major events that have already occurred and that had attracted serious forecasting efforts before they occurred (e.g., futures markets, London bookies, Nate Silver’s 538 site).

**Week 6: Individual-level interventions for improving probability judgments.** This session will introduce students to the best-validated techniques for improving the "calibration" and "resolution" of their probability judgments—and then extend those techniques to improve group judgment.

**Reading Assignments** Lovallo, D. & Kahneman, D. (2003). *Delusions of success: How optimism undermines executives’ decisions*. Harvard Business Review.

Makridakis, S. et al. (2010). *Why forecasts fail. What to do instead*. MIT Sloan Management Review.

Mauboussin, M. (2012). *The success equation: Untangling success and luck in business, sports, and investing*. Harvard Business Review Press. Chapters 4-7.

**Team Exercise:** Focus will be both on correcting individual biases and on the capacity of teams to correct errors beyond the capacity of isolated individuals to correct.

**Week 7. Team-level interventions for improving probability judgments.**

**Readings:**

Hackman, J. R. (2009). *Why teams don’t work*. Harvard Business Review.

Heath, C., Larrick, R., and Klayman, J. (1998). *Cognitive repairs: How organizational practices can compensate for individual shortcomings*. Research in organizational behavior.

Sunstein, C., Hastie, R. *Wiser: Getting Beyond Groupthink to Make Groups Smarter*, “Chapter 6: Eight Ways to Reduce Failures”, pp. 103-124.

**Team Exercise:** Students will conduct a hidden profiles exercise developed by the CIA to test their ability to identify the high-threat individual.

**Week 8. Miracles of Aggregation: Algorithmic solutions to improving probabilistic forecasts.**

**Reading Assignments**

Bonabeau, E. (2009). *Decisions 2.0: The power of collective intelligence*. MIT Sloan Management Review.

Schreiber, E. (2011). *Prediction markets*. In Jon Elster (ed). Crowd Sourcing.

Surowiecki, J. (2004). *The wisdom of the crowd*. Chapters 1-2.

**Team Exercise:** Students will solve a series of puzzles and estimation exercises.

**Week 9: Optimal forecasting frontiers: When do our efforts to improve foresight hit diminishing marginal returns?**

**Reading Assignments**

Bueno de Mesquita, B. (2009). *The predictioneer's game*. New York: Random House. Introduction-Chapter 3 and Chapters 8-11.

Taleb, N. et al. (2009). *Six mistakes executives make in risk management*. Harvard Business Review.

**Team exercises:** Extend precision questioning to Taleb's Black Swan and anti-fragility arguments.

**Week 10: Scenario planning.** Everyone loves a good story—and some scenario planners can spin elaborate images of possible futures that beguile the gullible.

**Readings:**

Schoemaker, P. & Day, G. (in press). *Integrating organizational networks, weak signals, strategic radars and scenario planning*. Technological Forecasting and Social Change.

CIA Tradecraft manual chapter on indicators of change

Cynthia Grabo, Anticipating Surprise, pp. 25-35.

**Team Exercise:** Students will respond to a scenario generation challenge.

**Week 11: The leader's dilemma.** Forecasters are often (rightly) wary of speaking truth to power. This session will explore situations in which forecasters are pressured to distort their public probability assessments.

**Readings:**

Superforecasting, chapter 10 and 12

Ingrid Verheul, Martin Caree, "Overoptimism among founders: the role of information and motivation", pp. 3-21.

**Team Exercises:** Teams will wrestle with an ethical conundrum: do great leader sometimes need to be great dissemblers? When, if ever, is it justifiable for leaders to project more confidence than they really feel in order to “buck up the troops”?

### **Week 12: Trade-Offs in Efforts to Improve Foresight.**

#### **Reading Assignments**

Edmondson, A. (2011). *Strategies for learning from failure*. Harvard Business Review.

Schoemaker, P. & Tetlock, P. E. (2012). *Taboo Scenarios: How to think about the unthinkable*. California Management Review.

Chapter 11, Superforecasters

**Team Exercises:** Teams will explore implicit or explicit taboos on the types of possibilities they felt pressure not to consider in their team deliberations or at Penn in general. A team that cannot identify its taboo topics is a team at risk of being blindsided. Teams will then also explore the adaptive functions of shared taboos.

### **Week 13: Resolving Debates Over How to Balance Process and Outcome Accountability in a Rapidly Changing World.**

#### **Reading Assignments**

Kahneman, D. (2011). *Thinking: Fast and slow*. New York: Farrar, Straus and Giroux. Chapters 15 and 16.

Tetlock, P.E. & Mellers, B. (2011). The intelligent management of intelligence agencies: Beyond accountability ping-pong. *American Psychologist*.

**Team Exercises:** Teams will be asked to decide whether they would prefer to be judged based on the accuracy of their forecasts (“outcome” accountability) or on the basis of the quality of the rationales they offer for their forecasts (“process” accountability). How fair or unfair is it to be evaluated by either an outcome or process accountability regimen?

### **Week 14: Closing Summary—and Class Presentations.**