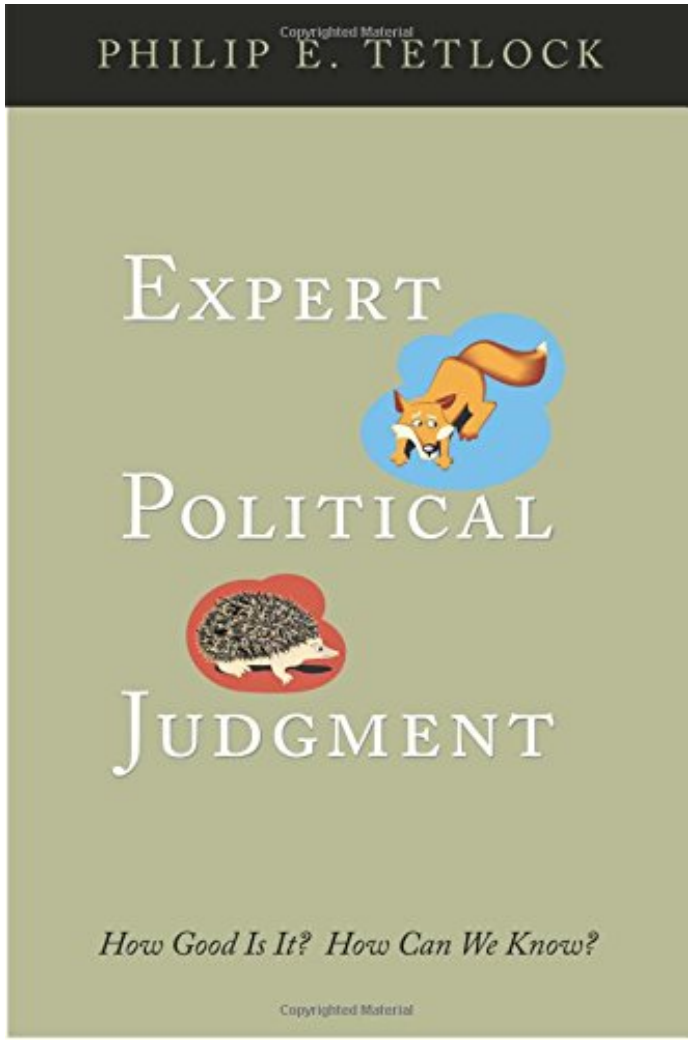
Pre-read for Tuesday, Sept 8: Predicting geopolitical events, part 2

Matthew J. Salganik

COS 597E/SOC 555 Limits to prediction

Fall 2020, Princeton University



20 years in the making

BOOKS DECEMBER 5, 2005 ISSUE

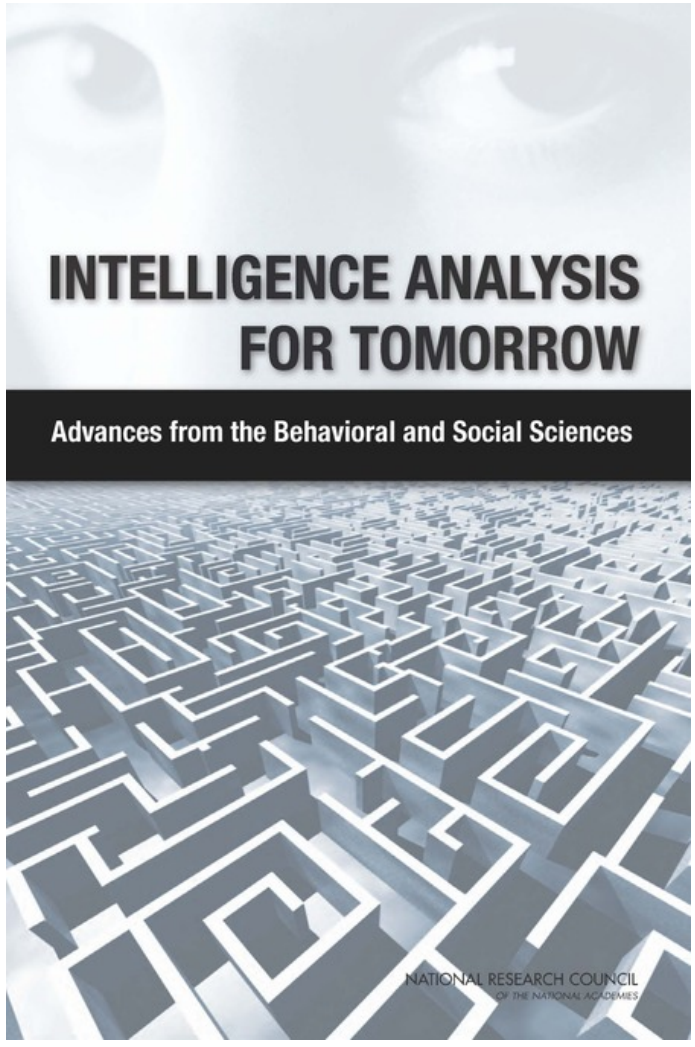
EVERYBODY'S AN EXPERT

Putting predictions to the test.

By Louis Menand

November 28, 2005

https://www.newyorker.com/magazine/2005/12/05/everybodys-an-expert



https://doi.org/10.17226/13040

Accuracy of forecasts in strategic intelligence

David R. Mandel[a,1] **and Alan Barnes**[b,2]

[a]Socio-Cognitive Systems Section, Defence Research and Development Canada, Toronto, ON, Canada M3K 2C9; and [b]Formerly of the Intelligence Assessment Secretariat, Privy Council Office, Ottawa, ON, Canada K1A OA3

https://doi.org/10.1073/pnas.1406138111

Intelligence Assessment Secretariat, part of the Canadian Security Intelligence Service

Predictions embedded in reports

- "The intense distrust that exists between Country X and Country Y is almost certain [9/10] to prevent the current relationship of convenience from evolving into a stronger alliance."
- "It is very unlikely [1/10] that either of these countries will make a strategic decision to launch an offensive in the coming six months."

- How do their predictions compare to the ones by Tetlock's experts?

- How do their predictions compare to the ones by Tetlock's experts?
- What variation is there in performance for different types of predictors or outcomes?

Reading notes:

- Tetlock (2005): roughly 15% of cases had ambiguous outcomes (p 296), Mandel and Barnes (2014): 20% of cases had ambiguous outcomes (p 10987)

Reading notes:

- Tetlock (2005): roughly 15% of cases had ambiguous outcomes (p 296), Mandel and Barnes (2014): 20% of cases had ambiguous outcomes (p 10987)
- Who picks tasks? It seems like analysts pick what they forecast (somewhat)

Reading notes:

- Tetlock (2005): roughly 15% of cases had ambiguous outcomes (p 296), Mandel and Barnes (2014): 20% of cases had ambiguous outcomes (p 10987)
- Who picks tasks? It seems like analysts pick what they forecast (somewhat)
- Hard or easy are coded by experts after the fact

From SI:
Low/moderate difficulty included judgments under most or all of the following conditions:
(i) availability of a substantial and credible information base,
(ii) involving a limited number of factors and/or largely a straight-line continuation of current trends,
(iii) little influence of irrational or unpredictable behavior, and
(iv) generally involving a short time horizon (several months).
Moderate/high difficulty included judgments affected by some of the following conditions:
(i) a limited and unreliable information base,
(ii) involving a wide range of complicated factors with multiple potential outcomes,
(iii) high likelihood of unpredictable behavior, and
(iv) involving a longer time horizon (a year or more).
There were 675 easier forecasts and 839 harder forecasts.
Providing some indication of reliability, there was less outcome variance in the easier set (VI = 0.19) than in the harder set (VI = 0.25). (But no measure of inter-coder reliability).

Judging political judgment

Philip Tetlock[1] and Barbara Mellers

Department of Psychology and The Wharton School, University of Pennsylvania, Philadelphia, PA 19104

https://doi.org/10.1073/pnas.1412524111

Reading notes

- How do they explain any difference in findings between Mandel and Barnes and Teltock (1995) and Meller and Tetlock (what comes after Tetlock 1995)?
- What does this mean for our ability to learn anything from an empirical study like this?

Danto

Cables

- How accurately can the perceived contemporaneous impression of importance judgements predict subsequent judgement by historians?
- How well can an ML model training on the cables predict subsequent judgement by historians?
- What do the errors of the humans and the algorithms teach us?

What to read next:

- "Bobby W." (2019) The Limits of Prediction—or How I Learned to Stop Worrying About Black Swans and Love Analysis "The key struggle for intelligence analysts is that they are made to produce and what their consumers think they can produce are often two different things."
- Doran (1999) https://www.jstor.org/stable/3186379Why Forecasts Fail: The Limits and Potential of Forecasting in International Relations and Economics
- What to read next from last class