

Superforecasting: the art and science of prediction

*By Philip Tetlock
and Dan Gardner
Random House Books
05 pp.*

*ISBN: 0: 84794758, 978-84794754
Review DOI
10.1108/FS-12-2016-0061*

This book continues the work of Philip Tetlock and collaborators on the accuracy of forecasters (mostly covering political and economic events) by reporting on a four-year tournament sponsored by the Intelligence Advanced Research Projects Activity (IARPA), an umbrella association of US intelligence agencies. One of the motivations for this experiment using real-world forecasting problems was the pressing need of policy-makers for good forecasts. Estimates indicate that the US intelligence community spends several billion dollars annually on geopolitical forecasts. Another reason lies in pessimistic findings on the accuracy of expert political forecasts reported in earlier studies. Avoiding many potential statistical pitfalls, Tetlock and his team identify a fraction of 2 per cent (from a total of 2,800) of volunteers who forecast at least 30 per cent better than the average and do so consistently over time. The book explains in detail the structure of the forecasting contest and identifies important psychological characteristics of these superforecasters.

Given the format of the study, the forecast tasks had to be quite short term. Only in this way was it possible to annually assess the accuracy of forecasts. Exemplary questions were of the following type: Will the prime minister of Italy (at the time Berlusconi) be forced out of office before the end of the year? Will the dollar against the euro pass a certain threshold level over the

course of the year? Participants in the study were recruited via professional societies, alumni associations, science blogs etc. The reward offered was a \$250 gift certificate at the end of every year for every contestant tackling a minimum of 25 problems a year. Furthermore, successful forecasters could count on being identified on a listing of the most successful contestants.

Most of the forecasting tasks were binary in nature (will something occur or not?). The forecast would thus come in the form of a yes or a no. Furthermore, forecasters were asked to report the probability with which they expected an outcome to occur. Hence, a forecaster who is totally sure of the answer would respond with yes (with probability 1) and no (with probability 0). A more cautious forecaster might likewise judge the yes-outcome as the more likely scenario. However, she might only attribute a probability of 0.7 to the yes-outcome, hence considering it plausible that the no-outcome can occur with a probability of 0.3. The particular criterion used in the study for the ranking of forecasters, the Brier score, takes this dimension of confidence into account. It has the feature of penalizing overconfidence. The higher the probability given to a scenario that does not materialize the worse is the score. Thus, the score does not simply judge the frequency of correct calls but it also judges a forecaster's calibration, i.e. the ability to accurately assess one's forecasting skill. If I know that my forecast is correct 60 per cent of the time for a specific type of task this is very valid information. It is relevant for anyone considering relying on my forecast because it addresses the reliability of the forecast and will influence the stakes taken in risky decisions.

What are typical attributes then of the group of people identified as superforecasters? This sort of individual tends to hold a non-deterministic worldview, i.e. these people accept that nothing is for certain. Such individuals tend to think in alternative scenarios and seek information to evaluate the likelihood. High performing forecasters further are what Tetlock and Gardner call actively open minded. This means that they continuously question their own views and tend to incorporate new information as it arrives. These forecasters show a need for cognition which shows in their stated preferences to solve puzzles and deal with mental challenges. They are further able to accept and integrate diverse views, and they are good at avoiding cognitive and emotional biases. High-performance forecasters are good at grading uncertainty which correlates with their ease with numbers. This further shows in their frequent yet gradual updating of forecasts as new information arrives. Many aspects thus emerging in the description of superior forecasting behavior address the relevance of the dichotomy of cognition suggested by [Kahneman \(2003\)](#). The so-called system 1 of human data processing (fast, automatic but prone to bias) and system 2 (deliberate, reflective but slow) tend to complement each other in our daily struggle for survival. Forecasters who perform systematically above average stand out in keeping system 2 activated continuously.

The authors of the book also identify – often with vivid descriptions of exemplary participants in the study – how well forecasters structure their task. A typical way of approaching a forecasting problem starts with the search for a base rate, i.e. an assessment of the frequency of occurrence of events similar to the one to be assessed. An example

would be the forecast of whether the recurrence of a certain infectious disease will lead to fatalities before the end of the year. A base rate estimate for such a task can be computed by researching available information on the percentage of years in which this disease has caused fatalities. If this frequency is, say, 0.20 and the forecast is done at the end of the first quarter, then a reasonable base rate would be 0.15. From this starting point, a good forecaster looks for more information that is relevant to the question as, e.g. information from the World Health Organization containing early warning signals regarding possible outbreaks of this particular disease.

A significant part of the book deals with issues concerning the aggregation of information contained in individuals' forecasts. As a benchmark, the idea of the "wisdom of crowds" was used. Not surprisingly, it was verified that averaging over individual forecasts yields superior results. However, starting with the second year of the study, forecast averaging was put into competition with alternative ways of aggregating information and several ingenious forms of integrating individual judgments generated superior performances. For example, groups of highly successful forecasters who were motivated to share data and ideas before finalizing their predictions did outstandingly well. Of course, as with any other interesting experiment, there remain some unanswered questions. One issue concerns the tasks actually selected by individual forecasters among a longer list of problems. A related issue then is how people perform in a professional setting where they have to routinely assess certain and non-negotiable forecasting questions. And one more critical point: one mechanism that tends to decrease the score of ordinary forecasters is the well-known

general tendency for overconfidence. With that in mind, it would be worth to reconsider the specific wording of questions.

Readers interested in further details on the highly interesting research agenda covered in this book will find stimulating reports in already published journal articles. As an example, the text by [Mellers et al. \(2015\)](#) gives many additional insights in a condensed form. This article coauthored by 12 researchers also points to the organizational challenge of running such a complex study and filtering through the massive amount of data gathered.

References

Kahneman, D. (2003), "Maps of bounded rationality: psychology for behavioral

economics", *American Economic Review*, Vol. 9 No. 5, pp. 449-475.

Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., Chen, E., Baker, J., Hou, Y., Horowitz, M., Ungar, L. and Tetlock, P. (2015), "Identifying and cultivating superforecasters as a method of improving probabilistic predictions", *Perspectives on Psychological Science*, Vol. 10 No. 3, pp. 267-281.

Tobias Rötheli

Tobias Rötheli is Professor at the Department of Economics, University of Erfurt, Erfurt, Germany.

About the author

The research of Tobias Rötheli focuses on the behavioral economics of expectations formation and business forecasting.

Tobias Rötheli can be contacted at: tobias.roetheli@uni-erfurt.de