

GRAY MATTER

The Power of Precise Predictions

By Philip E. Tetlock and J. Peter Scoblic

Oct. 2, 2015

IS there a solution to this country's polarized politics?

Consider the debate over the nuclear deal with Iran, which was one of the nastiest foreign policy fights in recent memory. There was apocalyptic rhetoric, multimillion-dollar lobbying on both sides and a near-party-line Senate vote. But in another respect, the dispute was hardly unique: Like all policy debates, it was, at its core, a contest between competing predictions.

Opponents of the deal predicted that the agreement would not prevent Iran from getting the bomb, would put Israel at greater risk and would further destabilize the region. The deal's supporters forecast that it would stop (or at least delay) Iran from fielding a nuclear weapon, would increase security for the United States and Israel and would underscore American leadership.

The problem with such predictions is that it is difficult to square them with objective reality. Why? Because few of them are specific enough to be testable. Key terms are left vague and undefined. (What exactly does "underscore leadership" mean?) Hedge words like "might" or "could" are deployed freely. And forecasts frequently fail to include precise dates or time frames. Even the most emphatic declarations — like former Vice President Dick Cheney's prediction that the deal "will lead to a nuclear-armed Iran" — can be too open-ended to disconfirm.

There is a familiar psychological mechanism at work here. One of us, Professor Tetlock, has been running lab studies since the early 1980s that show that if people expect that others will evaluate the accuracy of their judgments — that is, if people feel they will be held accountable for their views — then they tend to avoid cognitive pitfalls such as overconfidence and the failure to update beliefs in response to new evidence. Professor Tetlock and the psychologist Jennifer Lerner have demonstrated that accountability has this effect because it encourages people to pre-emptively think of ways in which they might be wrong — before others do it for them.

But when people make non-falsifiable predictions, they feel less accountable. After all, if a prediction can never be disproved, then it poses no reputational risk. That lack of accountability, in turn, encourages overconfidence and even more extreme predictions.

Non-falsifiable predictions thus undermine the quality of our discourse. They also impede our ability to improve policy, for if we can never judge whether a prediction is good or bad, we can never discern which ways of thinking about a problem are best.

The solution is straightforward: Replace vague forecasts with testable predictions. Will the International Atomic Energy Agency report in December that Iran has adequately resolved concerns about the potential military dimensions of its nuclear program? Will Iran export or dilute its quantities of low-enriched uranium in excess of 300 kilograms by the deal's "implementation day" early next year? Within the next six months, will any disputes over I.A.E.A. access to Iranian sites be referred to the Joint Commission for resolution?

Such questions don't precisely get at what we want to know — namely, will the deal make the United States and its allies safer? — but they are testable and relevant to the question of the Iranian threat. Most important, they introduce accountability into forecasting. And that, it turns out, can depolarize debate.

In recent years, Professor Tetlock and collaborators have observed this depolarizing effect when conducting forecasting "tournaments" designed to identify what separates good forecasters from the rest of us. In these tournaments, run at the behest of the Intelligence Advanced Research Projects Activity (which supports research relevant to intelligence agencies), thousands of forecasters competed to answer roughly 500 questions on various national security topics, from the movement of Syrian refugees to the stability of the eurozone.

The tournaments identified a small group of people, the top 2 percent, who generated forecasts that, when averaged, beat the average of the crowd by well over 50 percent in each of the tournament's four years. How did they do it? Like the rest of us, these "superforecasters" have political views, often strong ones. But they learned to seriously consider the possibility that they might be wrong.

What made such learning possible was the presence of accountability in the tournament: Forecasters were able to see their competitors' predictions, and that transparency reduced overconfidence and the instinct to make bold, ideologically driven predictions. If you can't hide behind weasel words like "could" or "might," you start constructing your predictions carefully. This makes sense: Modest forecasts are more likely to be correct than bold ones — and no one wants to look stupid.

This suggests a way to improve real-world discussion. Suppose, during the next ideologically charged policy debate, that we held a public forecasting tournament in which representatives from both sides had to make concrete predictions. (We are currently sponsoring such a tournament on the Iran deal.) Based on what we have seen in previous tournaments, this exercise would decrease

the distance between the two camps. And because it would be possible to determine a “winner,” it would help us learn whether the conservative or liberal assessment of the issue was more accurate.

Either way, we would begin to emerge from our dark age of political polarization.

Philip E. Tetlock, a professor at the University of Pennsylvania, is a co-author (with Dan Gardner) of “Superforecasting: The Art and Science of Prediction.” J. Peter Scoblic is a fellow with the International Security Program at New America and a doctoral student at Harvard Business School.

A version of this article appears in print on , Section SR, Page 10 of the New York edition with the headline: The Power of Precise Predictions