

# A Better Crystal Ball

## The Right Way to Think About the Future

By [J. Peter Scoblic and Philip E. Tetlock](#) (FOREIGN AFFAIRS, Nov-Dec 2020)

[November/December 2020](#)



John W. Tomac

Every policy is a prediction. Tax cuts will boost the economy. Sanctions will slow Iran's nuclear program. Travel bans will limit the spread of COVID-19. These claims all posit a causal relationship between means and ends. Regardless of party, ideology, or motive, no policymaker wants his or her recommended course of action to produce unanticipated consequences. This makes every policymaker a forecaster. But forecasting is difficult, particularly when it comes to geopolitics—a domain in which the rules of the game are poorly understood, information is invariably incomplete, and expertise often confers surprisingly little advantage in predicting future events.

These challenges present practical problems for decision-makers in the U.S. government. On the one hand, the limits of imagination create blind spots that policymakers tend to fill in with past experience. They often assume that tomorrow's dangers will look like yesterday's, retaining the same mental map even as the territory around them changes dramatically. On the other hand, if policymakers addressed all imaginable threats, the United States would need so large and expensive a national security establishment that the country could do little else. By many measures, it is nearing this point already. The United States has military bases in [more than 70 countries](#) and territories, boasts more than four million federal employees with security clearances, and fields 1.3 million active-duty troops, with another million in

reserve. According to one estimate, the United States [spends \\$1.25 trillion](#) annually on national security. When it comes to anticipating the future, then, the United States is getting the worst of both worlds. It spends untold sums of money preparing yet still finds itself the victim of surprise—fundamentally ill equipped for defining events, such as the emergence of COVID-19.

There is a better way, one that would allow the United States to make decisions based not on simplistic extrapolations of the past but on smart estimates of the future. It involves reconciling two approaches often seen to be at philosophical loggerheads: scenario planning and probabilistic forecasting. Each approach has a fundamentally different assumption about the future. Scenario planners maintain that there are so many possible futures that one can imagine them only in terms of plausibility, not probability. By contrast, forecasters believe it is possible to calculate the odds of possible outcomes, thereby transforming amorphous uncertainty into quantifiable risk. Because each method has its strengths, the optimal approach is to combine them. This holistic method would provide policymakers with both a range of conceivable futures and regular updates as to which one is likely to emerge. For once, they could make shrewd bets about tomorrow, today.

### **PLANNING FOR UNCERTAINTY**

Although widely used in business today, the first element of this duo—scenario planning—grew out of post–World War II national security concerns, specifically the overwhelming uncertainty of the nuclear revolution. Previously, martial experience was thought to offer some guidance through the fog of war. Nuclear weapons, however, presented a novel problem. With the newfound ability to destroy each other as functioning societies in a matter of minutes or hours, the United States and the Soviet Union faced an unprecedented situation. And unprecedented situations are, by definition, uncertain. They lack any analogy to the past that would allow decision-makers to calculate the odds of possible outcomes.

Still, early U.S. efforts at nuclear-war planning sought to turn that problem into a calculable one. During World War II, the Allies had great success with the new field of [operations research](#), the application of statistical methods to improve the outcome of tactical engagements. After the war, the RAND Corporation—a “think factory” that the U.S. Air Force established as a repository for leading researchers—hoped to parlay this success into a new, more rational approach to war, based less on the intuition of generals and more on the quantification afforded by models and data.

Unfortunately, methods that worked at the tactical level proved nearly farcical at the strategic level. As the historian David Jardini has chronicled, RAND’s first attempt to model a nuclear strategy ignored so many key variables that it nonsensically called for deploying a fleet of aging turboprop bombers that carried no bombs because the United States did not have enough fissile material to arm them; the goal was simply to overwhelm Soviet air defenses, with no regard for the lives of the pilots. In the wake of such failures, it became clear that analysts could not entirely banish uncertainty. In 1960, even Charles Hitch, a man predisposed to calculation by dint of being RAND’s top economist and president of what was then the Operations Research Society of America, cautioned, “No other characteristic of decision-making is as pervasive as uncertainty.”

**When it comes to anticipating the future, the United States is getting the worst of both worlds.**

That, of course, raised the question of how to formulate sensible strategy. Unexpectedly, it was a RAND mathematician and physicist, Herman Kahn, who [offered an answer](#). If the lived past could not shape strategy, perhaps the imagined future could. Frustrated with RAND's attempts to scientize war, Kahn devoted himself to crafting scenarios in the pursuit of "ersatz experience" that would prepare the United States for the future through what were essentially thought experiments. Policymakers could use these scenarios as "artificial 'case histories' and 'historical anecdotes,'" Kahn wrote, thus making up for a lack of actual examples or meaningful data. They would provide analogies where there were none.

Early methods of generating scenarios were often freewheeling and discursive. But after scenario planning migrated [to the business world](#), it took on more structured forms. The most recognizable is a two-by-two matrix in which planners identify two critical uncertainties and, taking the extreme values of each, construct four possible future worlds. Regardless of the specific shape they take, rigorous scenario-planning exercises all involve identifying key uncertainties and then imagining how different combinations could yield situations that are vastly different from what mere extrapolation of the present would suggest. By then "backcasting"—taking one of these imagined futures as a given and asking what conditions produced it—scenario planners derive both a story and a system. They come up with a plausible narrative of how a future happened and an internal logic that describes how it operates. Scenarios are not supposed to be predictive. They are meant to be provocative, challenging planners' assumptions, shaking up their mental models of how the world works, and giving them the cognitive flexibility to better sense, shape, and adapt to the emerging future.

The pandemic has occasioned a [renaissance in the use of scenarios](#), as organizations from think tanks to technology companies grapple with the question of what a "new normal" might look like and how soon it might arrive. But the national security community has long used scenarios to address some of its most wicked problems—particularly high-stakes issues that are in flux, such as the U.S.-Chinese relationship. This past summer, RAND released [a report](#) on Chinese grand strategy. It concluded with four scenarios that offered brief vignettes of China's possible place in the world 30 years from now. "Triumphant China" dominates the world stage in most domains, with a modern military and an innovative economy. "Ascendant China" is the preeminent power not only in Asia but in other regions, as well. "Stagnant China" has suffered from low growth and faces social unrest. And "Imploding China" experiences a crisis of existential proportions, in which domestic instability undercuts the country's international influence.

Although comprehensive, the wide range of these scenarios highlights the chief challenge of the method: If China's potential futures encompass rise, fall, and everything in between, how can they aid in the formulation of strategy and policy? Although this cornucopia of scenarios could lead policymakers to develop strategies that would improve the United States' position no matter which future comes to pass, in practice, having too many different versions of the future can make it nearly impossible to act. Good scenario planning puts boundaries on the future, but those limits are often not enough for decision-makers to work with. They need to know which future is most likely.

## **TURNING UNCERTAINTY INTO RISK**

Probabilistic forecasting—the second element of the duo—tries to address that shortcoming. Forecasters see scenario planning as maddeningly vague or, worse, dangerously misleading.

They not only point to the lack of consistent evidence to support the alleged benefits of scenario planning; they also argue that the compelling nature of a good story can trigger a host of biases. Such biases fuel irrationality, in part by tricking decision-makers into making basic statistical errors. For example, even though a detailed narrative may seem more plausible than a sparse one, every contingent event decreases the likelihood that a given scenario will actually transpire. Nevertheless, people frequently confuse plausibility for probability, assigning greater likelihood to specific stories that have the ring of truth. They might, illogically, consider a war with China triggered by a clash in the Taiwan Strait more likely than a war with China triggered by any possible cause.

In contrast to scenario planning's emphasis on imagination, forecasting tends to rely on calculation. Deductive approaches use models or laws that describe the behavior of a system to predict its future state, much like Newtonian mechanics allows astronomers to anticipate the position of the planets. Inductive approaches do not require such understanding, merely enough data and the assumption that the future will in some way reflect the past. This is how Netflix anticipates what you might like to watch or Amazon what you might want to buy, based purely on your previous actions. Increasingly, thanks to advances in artificial intelligence and machine learning, analysts use hybrid approaches. Meteorology is a good example, in which researchers combine sophisticated models and big data collection, which feed into each other and enable ever-better weather forecasts.

International politics poses a challenge for these methods because the laws governing the system are elusive or highly debatable, relevant data points are often unavailable or unprecedented, and thousands of variables interact in countless ways. History functions as a series of unfolding events, with highly contingent branching paths sometimes separated by mere happenstance. Tectonic shifts can hinge on seemingly mundane occurrences. That makes it hard to deduce future events from theoretical principles or to induce them from past experience.



Exhibiting military equipment in Changchun, China, August 201 Stringer / Reuters

As a result, historians and foreign policy experts are often bad forecasters. In 2005, one of us, Philip Tetlock, [published a study](#) demonstrating that seasoned political experts had trouble outperforming “dart-tossing chimpanzees”—random guesses—when it came to predicting global events. The experts fared even worse against amateur news junkies. Overconfidence was the norm, not the exception. When experts expressed 100 percent certainty that events would occur, those events materialized only 80 percent of the time. Yet there were pockets of excellence amid this unimpressive performance. Those who were surest that they understood the forces driving the political system (“hedgehogs,” in the philosopher Isaiah Berlin’s terminology) fared significantly worse than their humbler colleagues, who did not shy from complexity, approaching problems with greater curiosity and open-mindedness (“foxes”).

This distinction caught the eye of the Intelligence Advanced Research Projects Activity, which set up a geopolitical [“forecasting tournament,”](#) in which Tetlock participated. He recruited a team of volunteers to provide probabilistic answers to sharply defined questions, such as “Will the euro fall below \$1.20 in the next year?” or “Will the president of Tunisia flee to exile in the next six months?” By measuring the difference between estimates and the actual occurrence of events, Tetlock and his colleagues could calculate a score showing how “well-calibrated” the expectations of any given forecaster were with reality. By analyzing these data, Tetlock [discovered that the key](#) to more accurate geopolitical forecasting was to take people who were naturally numerate and open-minded, train them to think probabilistically and avoid common biases, and then group them so they could leverage the “wisdom of the crowd.” The best forecasters would approach seemingly intractable questions by decomposing them into parts, researching the past frequency of similar (if not precisely analogous) events, adjusting the odds based the uniqueness of the situation, and continually updating their estimates as new information emerged. By the end of the tournament, Tetlock’s top performers had achieved scores that were 30 percent better than those of career CIA analysts with access to classified information. Somehow, they had transmuted uncertainty into measurable risk.

The advantages of being able to put realistic odds on possible futures are obvious. It gives you a peek into the future. But even the best forecasters have their limits. If asked to predict events three to five years out, their performance becomes increasingly indistinguishable from random guessing. Still, many critical policy questions are short term: perhaps the most famous recent example concerned whether Osama bin Laden was in the Abbottabad compound in May 2011. Highly consequential short-term questions now include when a COVID-19 vaccine will be widely available. As of this writing, the smart money (68 percent probability) is on or before March 31, 2021.

**The greatest barrier to a clearer vision of the future is not philosophical but organizational.**

But to the extent that leaders need to make consequential, difficult-to-reverse decisions that will play out over the long run—the strategic choices that will give the United States an advantage over time—it becomes more difficult to link forecasts to policymaking. Well-calibrated forecasters, for instance, can estimate the likelihood that a skirmish with the Chinese navy in the South China Sea will result in at least two American deaths by December 31. But what policymakers really want to know is the extent to which China will threaten U.S. interests in the coming years and decades.

Answers to that type of inquiry are beyond the reach of forecasters because it is impossible to define precisely what constitutes an interest or a threat. To provide forecasts, questions must pass the “clairvoyance test,” which is to say that were it possible to pose the question to a

genuine clairvoyant, that omniscient seer must be able to answer it without having to ask for clarification. “Will I fall in love?” is not a forecasting question. “Will I marry Jane Smith by this time next year?” is.

From a policy perspective, then, the greatest challenge to forecasts is that although they can clarify slices of the future, they do not necessarily provide enough information to inform decision-making. Indeed, making a decision based on one specific forecast would be a mistake: the estimated probability of an event is a poor proxy for the significance of that event. “Will Vladimir Putin relinquish power within the next two years?” is a far different question from “What would Vladimir Putin’s abdication of power mean for U.S.-Russian relations?” The problem with forecasting is thus the exact opposite of the problem with scenarios: if the latter often provide too panoramic a view of the future to be useful, the former provides too narrow a glimpse.

### **AN ANSWER FOR THE FUTURE**

How should these different approaches to anticipating the future be linked? The answer lies in developing clusters of questions that give early, forecastable indications of which envisioned future is likely to emerge, thus allowing policymakers to place smarter bets sooner. Instead of evaluating the likelihood of a long-term scenario as a whole, question clusters allow analysts to break down potential futures into a series of clear and forecastable signposts that are observable in the short run. Questions should be chosen not only for their individual diagnostic value but also for their diversity as a set, so that each cluster provides the greatest amount of information about which imagined future is emerging—or which elements of which envisioned futures are emerging. As a result, the seductiveness of a particular narrative will not tempt decision-makers into mistaking plausibility for probability. Instead, preliminary answers to specific questions can provide a simple metric for judging in advance how the future is most likely to unfold—a metric that analysts can then refine once the event in question takes place or not.

Consider the scenarios RAND produced as part of its analysis of China’s grand strategy. The four scenarios envisioned for 2050 —“Triumphant China,” “Ascendant China,” “Stagnant China,” and “Imploding China”—can be roughly placed on a classic two-by-two matrix, with the strength of China’s political leadership on one axis and the strength of China’s economy on the other. A cluster of questions that would give a heads-up that history is on a “Triumphant China” trajectory might include “On December 31, 2020, will China exercise [de facto control](#) over Itu Aba (or Taiping Island) in the South China Sea (which is currently under the de facto control of Taiwan)?” “Will China’s GDP growth in 2023 exceed ten percent?” and “Among African audiences, when will the China Global Television Network have a higher weekly viewership than Voice of America?”

These questions are useful both individually and collectively. Knowing that top forecasters see an increased chance of China controlling the island (from, say, a ten percent probability to a 20 percent probability), for instance, would provide immediate tactical value to the U.S. Navy. It should not necessarily tip the balance in the debate over whether China will be “triumphant,” but if all the forecasts resulting from the question cluster are trending in the same direction, the United States may want to recalibrate its strategy. As forecasts change and individual questions are answered by the course of events, the view of the far-off future becomes a little bit clearer. Analysts can then update their scenarios and generate new clusters of questions.

They can thus develop a continually evolving sense of plausible futures, as well as a probabilistic estimate of which policies will yield the most bang for the buck today.



Taiwanese ships in the South China Sea, November 2016

J.R. Wu / Reuters

This method resembles the U.S. defense and intelligence community's use of indications and warnings. In the early 1960s, for example, the National Intelligence Council developed a list of actions—large troop maneuvers, for instance—that might precede an attack by the Sino-Soviet bloc. The idea was that tactical changes might provide an early warning of future strategic shifts. Indications and warnings have come to play an important role in many national security scenarios. Unfortunately, there are potential problems with scouring today's environment for hints of tomorrow. For one thing, as psychological research shows, having envisioned a particular scenario, humans are not only inclined to consider it more likely; they are also more prone to see evidence of its emergence—a form of confirmation bias that U.S. intelligence has battled for decades. For another, analysts are not particularly good at discerning in real time which events matter—which signposts are actually indicative of a particular future. Developments initially considered to be earthshattering may turn out to be significantly less important, whereas a story buried well beneath the day's headlines can end up changing the course of history. In a statistical analysis of nearly two million State Department cables sent in

the 1970s, for instance, one recent study demonstrated that U.S. diplomats were often bad at estimating the historical importance of contemporaneous events.

Linking scenarios to clusters of forecasting questions mitigates these problems. First, because the questions must be precise enough to pass the clairvoyance test, there is no wiggle room about what constitutes, say, large troop maneuvers. Second, because questions that disprove hypotheses often yield the greatest information, selecting questions for their diagnostic value decreases forecasters' susceptibility to confirmation bias. Third, much as diversified stock portfolios spread risk through multiple, uncorrelated investments, the diversity of question clusters prevents forecasters from overweighting a potentially unimportant signpost and mistakenly concluding that a particular scenario is coming to pass. Finally, and most important, because question clusters yield forecasts, one can attach meaningful probabilities to the likelihood that particular events will occur in the future. This provides a sort of advance early warning system. An event does not need to actually transpire for the United States to have actionable information. That, more than anything else, gives question clusters an advantage over traditional indications and warnings.

### **PLANNING IN PRACTICE**

To be useful, any vision of the future must be connected to decisions in the present. Scholars and practitioners often claim that scenario planning and probabilistic forecasting are incompatible given their different assumptions and goals. In fact, they mesh well. A scenario planner's conviction that the future is uncertain need not clash with a forecaster's quest to translate uncertainty into risk. Rather, the challenge lies in understanding the limits of each method. Question clusters make it possible to leverage the strength of each approach, transforming the abstract long term into the concrete short term so that leaders can understand the future quickly and act to stave off danger, seize opportunity, and strengthen national security.

The greatest barrier to a clearer vision of the future is not philosophical but organizational: the potential of combining scenario planning with probabilistic forecasting means nothing if it is not implemented. On occasion, the intelligence community has used forecasting tournaments to inform its estimates, but that is only a first step. Policymakers and consumers of intelligence are the ones who must understand the importance of forecasts and incorporate them into their decisions. Too often, operational demands—the daily business of organizations, from weighty decisions to the mundane—fix attention on the current moment.

Overcoming the tyranny of the present requires high-level action and broad, sustained effort. Leaders across the U.S. government must cultivate the cognitive habits of top forecasters throughout their organizations, while also institutionalizing the imaginative processes of scenario planners. The country's prosperity, its security, and, ultimately, its power all depend on policymakers' ability to envision long-term futures, anticipate short-term developments, and use both projections to inform everything from the budget to grand strategy. Giving the future short shrift only shortchanges the United States.

- J. PETER SCOBLIC is Co-Founder of Event Horizon Strategies, a Senior Fellow in the International Security Program at New America, and a Fellow at Harvard's Kennedy School.



- PHILIP E. TETLOCK is Leonore Annenberg University Professor at the University of Pennsylvania, Co-Founder of Good Judgment, and a co-author of *Superforecasting: The Art and Science of Prediction*.