Nurturing a Revolution in Decision-Making Through Improved Geopolitical Foresight

First Place Award
2016 USNI Naval Intelligence Essay Contest
by
LT Andreas Xenachis

Introduction

Today's national security policymakers face an operating environment far more challenging to fully comprehend than in previous generations. Threats and opportunities are more diffuse and diverse. Technological change has increased the pace of innovation and distributed access to both lethality and political power across a far larger set of actors. States and non-state groups alike have sought to challenge the U.S. not only through direct conflict, but by using indirect, hybrid approaches that blur boundaries and seek to exploit policymaking seams. Such challenges require new decision-making approaches. Diffuse threats can only be effectively addressed through institutional flexibility. Innovation must be met with agile responses. Hybrid conflict can only be challenged by a far more integrative approach to utilizing national power.

Underlying such interrelated challenges is the question of foresight: the ability to precisely characterize the operating environment, understand risks to the status quo, and accurately predict future changes to that environment. If individual decision-makers and institutional processes are unable to generate effective foresight, they will be far less likely to design successful policies that address the critical national security challenges facing the U.S. and its allies.

Foresight is central to the vision for the Navy's future promulgated by Admiral John Richardson, the Chief of Naval Operations (CNO), in February 2016. In "A Design for Maintaining Maritime Superiority," Admiral Richardson builds a strategy along four lines of effort—comprehensive sea power, improved learning, team-building, and partnerships—none of which can be achieved without improved foresight.[1] Maritime superiority is no longer simply about the physical domains of sea, air, space, and subsurface. The CNO mentions the "information domain" and the "'informationalized' environment" 17 times in the document's eight pages, more than any other

[1] Admiral John Richardson, "A Design for Maintaining Maritime Superiority," Online: http://www.navy.mil/ah_online/MaritimeSuperiority/index.html. See also John Richardson, "A Contest for Maritime Superiority, *Proceedings*, February 2016, pp. 24-27.

concept (including "sea" and "maritime").[2] The learning, team-building, and partnerships pillars are equally grounded in foresight and the information environment, as demonstrated by a brief sample of the document's key objectives and tasks:

- *"Apply the best concepts, techniques and technologies to accelerate learning as individuals, teams and organizations. Clearly know the objective and the theoretical limits of performance."*
- *"During execution, conduct routine and rigorous self-assessment. Adapt processes to be inherently receptive to innovation and creativity."*
- *"Expand the use of learning-centered technologies, simulators, online gaming, analytics and other tools as a means to bring in creativity, operational agility and insight."*
- *"Test and refine concepts through focused wargaming, modeling, and simulations."*
- *"Implement individual, team and organizational best practices to inculcate high velocity learning as a matter of routine."*
- *"Deepen the dialogue with private research and development labs, and academia."*[3]

In short, the CNO's design cannot be achieved without improving foresight to its utmost limits.

At the same time that the operating environment and our leaders are demanding more effective governmental foresight, some remarkable approaches to improved foresight are emerging from academia, business, and the public sector. This paper will focus on the most promising such example, which began as a pilot project sponsored by the U.S. Intelligence Advanced Research Projects Activity, as a case study to explain both the new state of the art and the institutional limitations preventing institutionalization of such innovations. It will ultimately argue that only determined, high-level leadership can lead to institutionalization of a revolutionary approach to governmental foresight that can sustain a new generation of national security policymaking successes.

---

2 Ibid.
3 Ibid.

Redefining the Frontier of Foresight

In 2010, the National Academy of Sciences (NAS) reported on the current scientific consensus underlying the fields of prediction and intelligence analysis.[4] A key conclusion was the need to "systematically track the accuracy of probabilistic forecasts that analysts routinely...make"—something which many outsiders would have (inaccurately) assumed was already being done inside the Intelligence Community.[5] Building on the NAS study, "the Intelligence Advanced Research Projects Activity (IARPA), the research and development branch of the Office of the Director of National Intelligence, launched a large-scale forecasting tournament designed to monitor the accuracy of probabilistic forecasts about events that occurred around the world."[6]

IARPA selected five academic groups through a robust proposal process and tasked them with forecasting real-world geopolitical outcomes on hundreds of questions selected by the government, using whatever methodology each group wished.[7] Some groups used computer algorithms, others emphasized expert knowledge, and others utilized variations on the "wisdom of crowds" (the insight that, for certain types of questions, simply averaging together a large number of individual predictions leads to fairly accurate results).[8] The teams knew they were competing against one another and against

---

4 B. Fischhoff and C. Chauvin (Eds.), *Intelligence analysis: Behavioral and social scientific foundations.* (Washington, DC: National Academies Press, 2011).

5 Barbara Mellers, Eric Stone, Pavel Atanasov, Nick Rohrbaugh, S. Emlen Metz, Lyle Ungar, Michael M. Bishop, Michael Horowitz, Ed Merkle, and Philip Tetlock, "The Psychology of Intelligence Analysis: Drivers of Prediction Accuracy in World Politics," *Journal of Experimental Psychology: Applied* (2015: Vol. 21, No. 1), pp. 1–14.

6 Ibid. For more on the IARPA Aggregative Contingent Estimate (ACE) program, see http://www.iarpa.gov/index.php/research-programs/ace

7 Examples of the types of questions that were asked included: "Will North Korea launch a new multistage missile before May 10, 2014? Will Russian armed forces enter Kharkiv, Ukraine, by May 10, 2014? Will there be a significant attack on Israeli territory before May 10, 2014? Will Robert Mugabe cease to be President of Zimbabwe by September 30, 2011? Will Greece remain a member of the EU through June 1, 2012?" (Steven Rieber, "Aggregative Contingent Estimation (ACE) Fact Sheet," Online: http://www.iarpa.gov/images/files/programs/ace/01-ACE.pdf)

8 The term "wisdom of crowds" was popularized by James Surowiecki's 2004 book of the same title. The concept itself dates back to the early 20th century: "In 1906 the legendary British scientist Sir Francis Galton went to a country fair and watched as hundreds of people individually guessed the weight [of an ox].…Their average guess—their collective judgment—was 1,197 pounds, one pound short of the correct answer." See Philip E.

a control group—but they did not initially know that they were also competing against intelligence analysts with access to classified information.[9]

A remarkable pattern of results emerged: a single team was consistently out-forecasting its competitors by shocking margins. In the first year alone, this team surpassed IARPA's end-of-project metric for forecasting accuracy improvement.[10] Even more impressively, according to *Washington Post* columnist David Ignatius, "the top forecasters...performed about 30 percent better than the average for intelligence community analysts who could read intercepts and other secret data."[11] The results were so stark that, after the competition's first two years, IARPA canceled the contracts with all other competitors to focus on the one outlier team.[12]

What set apart the one team that had such remarkable results? One critical factor was that one of its lead investigators, Professor Philip Tetlock of the Wharton School of Business at the University of Pennsylvania, started the project as a longtime skeptic of the possibility for significant improvement in forecasting accuracy. His earlier work had shown that so-called political experts performed no better than flipping a coin when predicting world events.[13] While all other groups were wedded to their forecasting approaches or enamored of particular technological solutions, Prof. Tetlock's team—The Good Judgment Project (GJP)—adopted a much more flexible methodology.[14] GJP recruited thousands

---

Tetlock and Dan Gardner, *Superforecasting: The Art and Science of Prediction* (New York: Crown, 2015), pp. 72-74.

9 David Ignatius, "More chatter than needed," *Washington Post*, November 1, 2013. Online: https://www.washingtonpost.com/opinions/david-ignatius-more-chatter-than-needed/2013/11/01/1194a984-425a-11e3-a624-41d661b0bb78_story.html

10 IARPA set the project's hoped-for improvement metrics as: 20% improvement over a control group in Year 1, 30% in Year 2, 40% in Year 3, and 50% at the end of the pilot project in Year 4. The outlier team—the Good Judgment Project—demonstrated 60% improvement in Year 1 and 78% improvement in Year 2. See Tetlock and Gardner, pp. 17-18 and 89-91.

11 David Ignatius, "More chatter than needed," *Washington Post*, November 1, 2013. Online: https://www.washingtonpost.com/opinions/david-ignatius-more-chatter-than-needed/2013/11/01/1194a984-425a-11e3-a624-41d661b0bb78_story.html

12 Tetlock and Gardner, pp. 17-18.

13 Philip Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, N.J.: Princeton University Press, 2005).

14 The Good Judgment Project was a collaboration between the Wharton School of Management at the University of Pennsylvania and the Haas School of Business at the University of California-Berkeley. See also Barbara

of individual forecasters and identified the consistent top performers ("Superforecasters")—then set out to distill what made them special. One key insight was that an open-minded temperament usually matters more than subject matter expertise in determining forecasting accuracy. Building the right kinds of teams and offering cognitive de-biasing training were also important.[15] But the factor that ultimately mattered most was practice. The most accurate forecasters treated foresight as a muscle that could be strengthened with repetition, but which would atrophy with disuse. By consistently training themselves to make precise forecasts and keeping score to create a tight feedback loop learning mechanism, good forecasters became great and average forecasters became above average.

The results were unambiguously impressive, not least to former skeptics. In an environment where process improvements resulting in a few percentage points of efficiency gain are celebrated, IARPA described the results—"a 50+% reduction in error compared to the current state-of-the-art"—as "the largest improvement in judgmental forecasting accuracy" ever observed in the public policy literature.[16] When asked in 2016 about what new analytic tool held the most promise for improving foresight, the Chairman of the National Intelligence Council singled out the GJP results.[17] Similar plaudits were heard from outside experts. In 2015, Harvard Professor and former senior White House administrator Cass Sunstein called the work "the most important scientific study I've ever read on prediction."[18] *New York Times* columnist David Brooks wrote in 2013 that if he were president or

---

Mellers, Eric Stone, Terry Murray, Angela Minster, Nick Rohrbaugh, Michael Bishop, Eva Chen, Joshua Baker, Yuan Hou, Michael Horowitz, Lyle Ungar, and Philip Tetlock, "Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions," *Perspectives on Psychological Science*, 2015: Vol. 10(3), pp. 267–281

15 Mellers et al, "Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions."

16 Steven Rieber, "Aggregative Contingent Estimation (ACE) Fact Sheet," Online: http://www.iarpa.gov/images/files/programs/ace/01-ACE.pdf. See also: http://www.iarpa.gov/index.php/research-programs/ace

17 Gregory Treverton, "Strategic Intelligence: A View from the National Intelligence Council," Center for Strategic and International Studies, March 4, 2016. Online: http://csis.org/event/strategic-intelligence-view-national-intelligence-council-nic

18 Cass R. Sunstein, "Prophets, Psychics and Phools: The Year in Behavioral Science," *Bloomberg View*, December 14, 2015. Online: http://www.bloombergview.com/articles/2015-12-14/prophets-psychics-and-phools-the-year-in-behavioral-scien

Secretary of State, he would want the predictions emerging out of the IARPA pilot project on his desk.[19]

Implications for the Intelligence Community

The foresight breakthroughs that emerged out of the Good Judgment Project bear directly on the central analytic mission of the Intelligence Community. However, intelligence leaders have not yet made any concerted effort to systematically incorporate such breakthroughs into the functioning of their organizations. To be sure, a few disparate elements throughout the intelligence bureaucracy—whether individual analysts, instructors, or managers of sub-organizations—have taken the initiative to incorporate these emerging insights into their own work.[20] But there is no evidence yet of any larger organizational decision to incorporate the training, talent management, and analytic benefits of such forecasting breakthroughs into the fabric of the intelligence bureaucracy. No champion for broadly institutionalizing such reform has yet publicly emerged, either within the Intelligence Community or among the oversight bodies in the legislative branch.[21]

Implementation of such breakthroughs is also hampered by bureaucratic inertia that has not kept up with recent advances. As Prof. Tetlock and Welton Chang write,

> "Current [intelligence] training is anchored in a mid-twentieth century understanding of psychology….encouraging a narrow perspective on the flaws of intuition and a correspondingly narrow search for remedies. The result has been a defensive mindset aimed at avoiding

---

ce. Sunstein served as administrator of the White House Office of Information and Regulatory Affairs during the Obama administration and is currently the Robert Walmsley university professor at Harvard Law School.

19 "If I were President Obama or John Kerry, I'd want [Good Judgment's] predictions on my desk." in David Brooks, "Forecasting Fox," *N.Y. Times*, March 22, 2013, p. A29. Online: http://www.nytimes.com/2013/03/22/opinion/brooks-forecasting-fox.html

20 Perhaps the most notable such move to date has been the commitment by the Chairman of the National Intelligence Council to utilize forecasts from the internal, classified Intelligence Community Prediction Market (ICPM) as a technique to "red team" and cross-check analytic insights. See Treverton, "Strategic Intelligence: A View from the National Intelligence Council," Center for Strategic and International Studies, March 4, 2016. Online: http://csis.org/event/strategic-intelligence-view-national-intelligence-council-nic

21 IARPA itself is an organization focused on sponsoring research and testing pilot projects, rather than institutionalizing successes when they arise.

> mistakes, not a proactive mindset aimed at getting it right by reducing uncertainty as aggressively as possible."[22]

Even more concerning is the lack of an evidence-based approach in evaluating existing analytic training, which "has never been validated against objective benchmarks of good judgment….For too long the intelligence community has shackled itself to a system of training that it never tested."[23] Such bureaucratic inertia has typically been broken only in the wake of a crisis, at greater expense and disruption than would otherwise have been the case. Even Sherman Kent, famed as the "father of intelligence analysis," was unable to sell a minor proposed change to regularly use percentages rather than vague verbiage (such as "somewhat likely") in intelligence assessments at the CIA. It took the debacle of the Iraq WMD assessments in 2003 to finally implement such a change in the face of organizational risk-aversion—four decades after Kent initially proposed it.[24] Intelligence policymakers should therefore seize the opportunity to begin carefully integrating these novel techniques now, before the next crisis.

Intelligence analytic training must begin to systematically track forecasting accuracy in a structured but realistic environment, such as the forecasting tournament in the IARPA pilot project. Such tracking should not be punitive, but rather wholly oriented toward individual and organizational foresight improvement. Training must focus on more than an adherence to an analytical style and methodological standard, but on better outcomes through an updated psychological understanding of the error-balancing and active open-mindedness required for maximally accurate forecasting. But even more than classroom training, leaders must demand realistic, continuous practice, which is the only way to create the tight feedback loop necessary for systematic foresight improvement. Such a regimen

---

22 Welton Chang and Philip E. Tetlock, "Rethinking the training of intelligence analysts," *Intelligence and National Security* (January 2016), pp. 1-18. DOI: 10.1080/02684527.2016.1147164. The article continues, "Current [intelligence] training is anchored in a mid-twentieth century understanding of psychology that focuses on checking over-confidence and rigidity but ignores the problems of under-confidence and excessive volatility."

23 Ibid.

24 See Tetlock and Gardner, pp. 53-57.

would lead to more effective individual analysts, but even more importantly, to improved organizational outcomes as intelligence organizations are able to provide policymakers with more accurate understandings of the operational environment and risk assessments.

Implications for National Security Policymaking

U.S. legislators, diplomats, aid administrators, and military and interagency planners require an improved foresight instinct as much as intelligence analysts do.[25] Since World War II, the greatest U.S. strategic advantage on the international stage has not been its impressive technology, but rather a national security decision-making process that has been able to operate inside an opponent's decision cycle.

Technology plays a key role in that process, but people are at its core. Unfortunately, many private- and public-sector decision makers today over-fixate and defer to purely engineering solutions—whether that means "big data" algorithms or some new "computer simulation." Government policy often favors hardware over organizational process improvements. It is often easier to convince the Department of Defense to buy a new machine than to try a new human organizational process. Institutionalizing new human approaches—even clear successes like those emerging from the IARPA forecasting pilot project—faces higher obstacles than pitching a new automated solution that promises to take the human out of the loop. But, while it plays an important role, technology *is not* synonymous with innovation, especially when it comes to the complexity of geopolitical foresight and risk assessment. Technology is merely a tool that enables innovation, which is a fundamentally human process of improvement. Some things cannot be fully automated, and that includes judgment under uncertainty. The military organization that has never lost sight of this fact is the special operations

[25] Interview with a member of the State Department Policy Planning Staff, January 27, 2016.

forces (SOF) community. Despite having access to some of the most incredible technologies ever imagined, the SOF community continues to adhere to the first of its "SOF Truths": "Humans are more important than Hardware."[26] Such mindsets make organizations such as U.S. Special Operations Command (SOCOM) one of the most promising candidates as early adopters of breakthroughs in foresight demonstrated over the past few years. It is also heartening to understand that new national security organizations, such as the Secretary of Defense's new Defense Innovation Board populated by many giants of science and Silicon Valley, have made it clear they will not lose sight of the human element in a milieu of whiz-bang technological solutions.[27]

With this human element at the core, new methodologies to improve individual and organizational foresight instinct should be broadly integrated throughout the bureaucracy. Polcymakers at every level must be trained to improve forecasting skills, rather than focusing solely on their subject matter expertise and hoping foresight ability will develop organically. Doctrinal publications should be updated to include the training techniques for improving forecasting.[28] Organizations such as the Navy Staff and the Joint Staff must institutionalize these techniques in their everyday functioning. New structured insights from the IARPA Good Judgment Project pilot should supplement existing techniques, such as wargaming and scenario planning, in helping policymakers better understand risk and trendlines. Foresight is too important to be left to any single element of national security policymaking—planners and operators have as much to gain as intelligence organizations from such new insights.

---

26 "SOF Truths," United States Special Operations Command. Online: http://www.soc.mil/USASOCHQ/SOFTruths.html

27 Interview with Defense Innovation Board staffer, July 27, 2016. For background, see Defense Media Activity, "Pentagon to Establish Defense Innovation Advisory Board," March 2, 2016. Online: http://www.defense.gov/-News-Article-View/Article/684366/pentagon-to-establish-defense-innovation-advisory-board

28 The most relevant publications would include Joint Publication (JP) 5-0, "Joint Operation Planning"; JP 2-0, "Joint Intelligence"; JP 2-01.3, "Joint Intelligence Preparation of the Operational Environment"; and the service-specific equivalents thereof; as well as the Joint Staff J7 "Insights and Best Practices Focus Papers" series. See http://dtic.mil/doctrine/

Conclusion

In one of the most renowned ancient commentaries on the Talmud, a famed rabbi explains the difference between Moses and other prophets of the Old Testament: "All the prophets had a vision of God as He appeared through nine [distorted] lenses," but "Moses saw God through one lens" alone.[29] Similar insights about mankind's imperfect perception but improvable foresight can be found in all major religious traditions.[30] While mankind is unlikely to ever cast aside all of the distorting filters that affect our perceptions, recent breakthroughs have shown that we can systematically improve our individual and organizational foresight ability. Foresight is a muscle that can be systematically strengthened. Taking advantage of such novel insights presents an important leadership opportunity for military and civilian decision-makers facing an uncertain and challenging operating environment.

---

[29] *Leviticus Rabbah* (5th Century CE), 1:14. The quoted passage is ascribed to Rabbi Judah bar Ilai (2nd century CE). A similar theme can be found in the Babylonian Talmud: "All the prophets gazed through a speculum that does not shine, while Moses our teacher gazed through a speculum that shines." (B.T. Yevamot 49B) See also, Abraham Joshua Heschel and Gordon Tucker (ed. and tr.), *Heavenly Torah: As Refracted Through the Generations* (New York: Continuum, 2006), p. 308.

[30] From the Christian tradition, see, for example, Paul's First Epistle to Timothy ("God...who lives in unapproachable light, whom no one has seen or can see," 1 Timothy 6:15-16) or John Milton's 1644 book *Areopagitica* ("For such is the order of God's enlightening his Church, to dispense and deal out by degrees his beam, so as our earthly eyes may best sustain it."). In the Hindu and Buddhist traditions, mankind's imperfect understanding of the world is captured through the concept of *maya*: "The world does not seem to be what it is....The world...is *maya*. The term *maya* has been translated as 'illusion,' but then it does not concern normal illusion. Here 'illusion' does not mean that the world is not real and simply a figment of the human imagination. 'Maya' means that the world is not as it seems; the world that one experiences is misleading as far as its true nature is concerned." (Hendrick M. Vroom and Lucy Jansen (tr.), *No Other Gods: Christian Belief in Dialogue with Buddhism, Hinduism, and Islam* (Grand Rapids, MI: Wm. B. Eerdmans Publishing, 1996), p. 57.)