

Creation of a National Institute for Analytic Methods

Toward Improving Intelligence Analysis

Steven Rieber and Neil Thomason

Traditionally, analysts at all levels devote little attention to improving how they think. To penetrate the heart and soul of the problem of improving analysis, it is necessary to better understand, influence, and guide the mental processes of analysts themselves.

— Richards J. Heuer, Jr.[1]

The opinions of experts regarding which methods work may be misleading or seriously wrong.

The United States needs to improve its capacity to deliver timely, accurate intelligence. Recent commission reports have made various proposals aimed at achieving this goal. These recommendations are based on many months of careful deliberation by highly experienced experts and are intuitively plausible. However, a considerable body of evidence from a wide range of fields indicates that the opinions of experts regarding which methods work may be misleading or seriously wrong. Better analysis requires independent scientific research. To carry out this research, the United States should establish a National Institute for Analytic Methods, analogous to the National Institutes of Health.

While much has been written about how to improve intelligence analysis, this article will show *how to improve the process of improving analysis*. The key is to conduct scientific research to determine what works and what does not, and then to ensure that the Intelligence Community uses the results of this research. [2]

Expert Opinions Can Be Unreliable

The reports of recent commissions examining the intelligence process—including the Senate Select Committee on Intelligence and the special presidential commission on Iraq weapons of mass destruction[3]—incorporate recommendations for improving analysis. These proposals, which include establishing a center for analyzing open-source intelligence and creating “mission managers” for specific intelligence problems, make intuitive sense.

We want to suggest, however, that this intuitive approach to improving intelligence analysis is insufficient. Examples from a wide range of fields show that experts’ opinions about which methods work are often dead wrong:

- For decades, steroids have been the standard treatment for head-injury patients. This treatment “makes sense” because head trauma results in swelling and steroids reduce swelling. However, a recent meta-analysis involving over 10,000 patients shows that giving steroids to head-injury patients apparently increases mortality.[4]
- Most police departments make identifications by showing an eyewitness six photos of possible suspects simultaneously. However, a series of experiments has demonstrated that presenting the photos sequentially, rather than simultaneously, substantially improves accuracy.[5]
- The nation’s most popular anti-drug program for school-age children, DARE (Drug Abuse Resistance Education), brings police officers into classrooms to teach about substance abuse and decisionmaking and to boost students’ self-esteem. But two randomized controlled trials involving nearly 9,000 students have shown that DARE has no significant effect on students’ use of cigarettes, alcohol, or illicit drugs.[6]
- Baseball scouting typically is done intuitively, using a traditional set of statistics such as batting average. Scouts and managers believe that they can ascertain a player’s potential by looking at the

statistics and watching him play. However, their intuitions are not very good and many of the common statistical measures are far from ideal. Sophisticated statistical analysis reveals that batting average is a substantially less accurate predictor of whether a batter will score than on-base percentage, which includes walks. The Oakland A's were the first team to use the new statistical techniques to dramatically improve their performance despite an annual budget far smaller than those of most other teams.[7]

These examples and many others illustrate two important points. First, even sincere, well-informed experts with many years of collective experience are often mistaken about what are the best methods. Second, the only way to determine whether the conventional wisdom is right is to conduct rigorous scientific studies using careful measurement and statistical analysis. Prior to the meta-analysis on the effects of steroids, there was no way of knowing that they were counterproductive for head injuries. And without randomized controlled studies, we would not have learned that DARE fails to reduce cigarette, drug, and alcohol use. Experts' intuitive beliefs about what works are not only frequently wrong, but also are generally not self-correcting.

Caution Advised

Consequently, we should be skeptical about the numerous recent proposals for improving intelligence analysis. The recommendations generally are based on years of experience, deep familiarity with the problems, careful reflection, and a sincere desire to help—all of which may lead to reforms that do as much harm as good. Some of the experts' sincere beliefs may be correct; others may be widely off the mark. Without systematic research, it is impossible to tell.

Some high-quality research relevant to intelligence analysis has already been done, but it is virtually unknown within the Intelligence Community. Consider, for example, devil's advocacy. Both the Senate report on Iraq's weapons of mass destruction (WMD) and the report of the president's commission proposed the use of devil's advocates.[8] In fact, devil's advocacy and "red teams"—which construct and press an alternate interpretation of how events might evolve or how information might be interpreted—are the only specific analytic techniques recommended by the Senate report, the president's commission report, and the 2004 Intelligence Reform Act.[9] None of these reports, however, mentions the

research on devil's advocacy, which is quite equivocal about whether this technique improves group judgment.[10] Some research suggests that devil's advocates may even aggravate groupthink (the tendency of group members to suppress their doubts).[11] As Charlan Nemeth writes:

. . . the results . . . showed a negative, unintended consequence of devil's advocate. The [devil's advocate] stimulated significantly more thoughts in support of the initial position. Thus subjects appeared to generate new ideas aimed at cognitive bolstering of their initial viewpoint but they did not generate thoughts regarding other positions. . . .[12]

Irving Janis, the author of *Groupthink*, suggested such possibilities over 30 years ago. Janis describes the use of devil's advocates by President Lyndon B. Johnson's administration:

[Stanford political scientist] Alexander George also comments that, paradoxically, the institutionalized devil's advocate, instead of stirring up much-needed turbulence among the members of a policy-making group, may create 'the comforting feeling that they have considered all sides of the issue and that the policy chosen has weathered challenges from within the decision-making circle.' He goes on to say that after the President has fostered the ritualized use of devil's advocates, the top-level officials may learn nothing more than how to enact their policy-making in such a way as to meet the informed public's expectation about how important decisions should be made and 'to project a favorable image into the instant histories that will be written shortly thereafter.'[13]

Thus, once institutionalized, the principal effect of devil's advocates may be to protect the Intelligence Community from future criticism and calls for reform. The scientific evidence shows that we cannot exclude the possibility that adopting the recommendations of the recent commission reports may be counterproductive.

Identifying What the Research Says

The first element in improving the process of improving analysis is to *find out what the existing scientific research says*. Not all of the existing research on how to improve human judgment is negative. Here are some promising results from this research:

- Argument mapping, a technique for visually displaying an argument's logical structure and evidence, substantially enhances critical thinking abilities.[14]
- Systematic feedback on accuracy makes judgments more accurate. [15]

- There are effective methods to help people easily avoid the omnipresent and serious fallacy of base-rate neglect.[16]
- Combining distinct forecasts by averaging usually raises accuracy, sometimes substantially.[17]
- Consulting a statistical model generally increases the accuracy of expert forecasts. [18]
- A certain cognitive style, marked by open-mindedness and skepticism toward grand theories, is associated with substantially better judgments about international affairs.[19]
- Simulated interactions (a type of structured role-playing) yields forecasts about conflict situations that are much more accurate than those produced by unaided judgments or by game theory.[20]

Applying the Research to Intelligence

While each of these findings is promising, almost none of this research has been conducted on analysts working on intelligence problems. Thus, the second element of improving the process of improving analysis is to *initiate systematic research on promising methods for improving analysis*. [21]

Each of the analytic methods mentioned above suggests numerous lines of research. In the case of argument mapping, for example, questions that should be investigated include: Do argument maps improve analytic judgment? In which domains (political, economic, military, long-range, or short-range forecasts) are argument maps most effective? How can analysts be encouraged to use the results of argument mapping in their written products? How can this method be effectively taught? If devil's advocates use argument maps, will their objections be taken more seriously?

The only reliable way to answer each of these questions is through scientific studies carefully designed to measure the relevant factors, control for extraneous influences, distinguish causation from correlation, and produce sizable effects. Intelligence analysts and other experts will certainly have opinions about how best to employ argument maps; in some cases, the experts may even agree with one another. But while the expert opinions should be considered in designing the research, they should not be the last word, since they may be mistaken.

Evaluation and development should be ongoing and concurrent and should provide feedback to the next round of evaluation and development,

in a spiraling process. Evaluation results will suggest ways of refining promising techniques, and the refined techniques can then be assessed.

Encouraging Use of New Methods

It is essential that there be serious research both inside and outside the analysis sector itself. American universities can become one of our great security assets. Techniques for improving analytic judgment can be tested initially on university students (both undergraduate and graduate); promising methods can then be refined and tested further by contractors, including former analysts, with security clearances. Techniques that are easy to employ and that substantially increase accuracy in these preliminary stages of evaluation could then be tested with practicing analysts.

It is essential to expose analysts only to methods that they are likely to use, and use well. Subjecting them to cumbersome or ineffective techniques would only waste their time and increase their possible skepticism about new methods.

Research should investigate not only which techniques improve analytic judgment, but also how to teach these techniques and how to get analysts to use them. Analytic methods that produce excellent results in the laboratory will be worthless if not used, and used correctly, by practicing analysts. Thus the third element of improving the process of improving analysis is to *conduct research on how to get promising analytic methods effectively taught and used.*

Communicating with Consumers

The purpose of intelligence analysis is to inform policymakers to help them make better decisions. Accuracy, relevance, and timeliness are not enough; intelligence analysis must effectively convey information to the consumer. No matter how cogently analysts reason, their work will fail in its purpose if it is not correctly understood by the consumer. Thus the fourth element of improving the process of improving analysis is to *conduct research on improving communication to policymakers.*

- How analysts should communicate their judgments to policymakers is yet another issue on which opinions are plentiful but systematic research is scarce. Some important questions here are:

- How can tacit assumptions be made explicit and clear? Can visual representations of reasoning, such as structured argumentation, usefully supplement prose and speech?
- How can the differences between analysts (or agencies) be communicated most effectively?
- What are the best ways for analysts to express judgments that disagree with the views of policymakers?
- Is the ubiquitous PowerPoint presentation a good way to present complex information? Or does it “dumb down” complex issues?[22]
- Forty years ago, Sherman Kent showed that different experts in international affairs had very different understandings of words like “probable” and “likely,” and that these differences produced serious miscommunication. How can this ongoing cause of miscommunication be alleviated?[23]

These questions can be systematically answered only through scientific research. Associated with each question is a cluster of research issues. Take, for instance, the question of how to communicate probability. Should analysts’ probabilistic judgments be conveyed verbally, numerically, or through a combination of the two? If verbal expressions are used, should they be given common meanings across analysts and agencies? Or should analysts assign their own numerical equivalents (making them explicit in their finished intelligence)? Should probabilistic statements be avoided altogether in favor of a discussion of possible outcomes and the reasons for each?

A National Institute

As shown by examples from other fields, systematic research can dramatically improve longstanding practices. This sort of research should be done on all aspects of intelligence analysis, including analytic methods, training, and communication to policymakers. To be most useful, the research should be well funded, coordinated, and held to the highest scientific standards. This requires an institutional structure. The National Institutes of Health provide an excellent model: NIH conducts its own research and funds research in medical centers and universities across the world.[24]

Just as NIH improves our nation’s health, a National Institute for Analytic Methods (NIAM) would enhance its security. To ensure that NIAM research would be of unimpeachable scientific caliber, it should work closely with,

but independently of, the Intelligence Community. In a similar vein, the president's WMD Commission recommends the establishment of one or more "sponsored research institutes":

We envision the establishment of at least one not-for-profit 'sponsored research institute' to serve as a critical window into outside expertise for the Intelligence Community. This sponsored research institute would be funded by the Intelligence Community, but would be largely independent of Community management.[25]

The Commission points out that "there must be outside thinking to challenge conventional wisdom, and this institute would provide both the distance from and the link to the Intelligence Community to provide a useful counterpoint to accepted views." [26] While the sponsored research institutes envisioned by the WMD Commission would tackle substantive issues, the NIAM would confront the equally important problems of developing, teaching, and promoting effective analytic methods.

To achieve this excellence and independence, a leadership team consisting of preeminent experts from inside and outside government is essential. Such a team is probably the only means to ensure that the research would be scientifically rigorous and adventurous, and that reform proposals would be truly evidence based. Many people mistakenly believe that they know how to do social-scientific research. However, this research is difficult, the methodology is complex and statistically sophisticated, and established results are often counter-intuitive. Only if guided by scientists of the highest caliber would evidence-based analytic methods advance as rapidly as their importance demands.

There are also political and bureaucratic reasons for having an expert leadership team. Without the prestige, influence, and financial clout of such a panel, bureaucratic inertia might prevent evidence-based reforms from being adopted. Bureaucratic rigidity is likely to become particularly serious as the intense political pressure for intelligence community reform diminishes. Initiating, funding, and coordinating research on all aspects of intelligence analysis is a large set of tasks. To perform these well, NIAM's budget would have to be adequate. When the Institute is fully running, a budget of 1–2 percent of NIH's may be appropriate.

A National Institute for Analytic Methods would contribute to long-term intelligence reforms in an unusual way. Most reforms become institutionalized and, thereafter, are rarely reevaluated until a subsequent crisis occurs. NIAM's evidence-based reforms would be very different.

Because science itself is a self-correcting process, NIAM-generated science would ensure that evidence-based reforms continue indefinitely. Thus, intelligence reforms would continue to improve analysts' effectiveness long after the current political urgency fades.

Footnotes

[1] *Psychology of Intelligence Analysis* (Washington: CIA Center for the Study of Intelligence, 1999), 173.

[2] Steven Rieber would like to express his deep appreciation to the Kent Center for Analytic Tradecraft for providing a stimulating environment for thought and discussion. The opinions expressed here are the authors' alone.

[3] *Report of the Senate Select Committee on Intelligence on the US Intelligence Community's Prewar Intelligence Assessments on Iraq*, 7 July 2004, and *The Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction Report to the President of the United States* [hereafter *WMD Commission Report*], 31 March 2005.

[4] CRASH Trial Collaborators, "Effect of Intravenous Corticosteroids on Death Within 14 days in 10008 Adults with Clinically Significant Head Injury (MRC CRASH Trial): Randomized Placebo-Controlled Trial," *Lancet* 364 (2004): 1321–28.

[5] Gary L. Wells and Elizabeth A. Olson, "Eyewitness Testimony," *Annual Review of Psychology* 54 (2003): 277–95. For an illuminating account of the damage caused by ongoing institutional resistance to evidence-based reform of eyewitness practices, see Atul Gawande. "Under Suspicion: The Fugitive Science of Criminal Justice," *New Yorker*, 8 January 2001: 50–53.

[6] Cheryl L. Perry, Kelli A. Komro, Sara Veblen-Mortenson, Linda M. Bosma, Kian Farbakhsh, Karen A. Munson, Melissa H. Stigler, and Leslie A. Lytle, "A Randomized Controlled Trial of the Middle and Junior High School D.A.R.E. and D.A.R.E. Plus Programs," *Archives of Pediatric and Adolescent Medicine* 157 (2003): 178–84.

[7] Michael Lewis, *Moneyball* (New York: Norton, 2003).

[8] *Report of the Senate Select Committee on Intelligence*, 21, and *WMD Commission Report*, 407.

[9] *House Report 108-796, Intelligence Reform and Terrorism Prevention Act of 2004*, Conference Report to Accompany 2845, 108th cong., 2nd sess., 35.

[10] Gary Katzenstein, "The Debate on Structured Debate: Toward a Unified Theory," *Organizational Behavior and Human Decision Processes* 66 (1996): 316–32; Alexander L. George and Eric K. Stern, "Harnessing Conflict in Foreign Policy Making: From Devil's to Multiple Advocacy," *Presidential Studies Quarterly* 32 (2002): 484–508.

[11] While it is certainly true that groups sometimes suppress their doubts, there is considerable debate over the mechanisms of such suppression. Many people mistakenly identify all suppression of doubts with groupthink: "The unconditional acceptance of the groupthink phenomenon without due regard for the body of scientific evidence surrounding it leads to unthinking conformity to a theoretical standpoint that may be invalid for the majority of circumstances." Marlene E. Turner and Anthony R. Pratkanis, "Twenty-Five Years of Groupthink Theory and Research: Lessons from the Evaluation of a Theory," *Organizational Behavior and Human Decision Processes* 73 (1998): 105–15

[12] Charlan Nemeth, Keith Brown, and John Rogers, "Devil's Advocate vs. Authentic Dissent: Stimulating Quantity and Quality," *European Journal of Social Psychology* 31 (2001): 707–20.

[13] Irving L. Janis, *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*, 2nd ed. (Boston: Houghton Mifflin, 1982), 268.

[14] Tim van Gelder, Melanie Bissett, and Geoff Cumming, "Cultivating Expertise in Informal Reasoning," *Canadian Journal of Experimental Psychology* 58 (2004): 142–52.

[15] Fergus Bolger and George Wright, "Assessing the Quality of Expert Judgment," *Decision Support Systems* 11 (1994): 1–24.

[16] Peter Sedlmeier and Gerd Gigerenzer, "Teaching Bayesian Reasoning in Less Than Two Hours," *Journal of Experimental Psychology: General* 130 (2001): 380–400.

[17] J. Scott Armstrong, "Combining Forecasts," in J. Scott Armstrong, ed., *Principles of Forecasting* (Boston, MA: Kluwer, 2001), 417–39.

- [18] William M. Grove, David H. Zald, Boyd S. Lebow, Beth E. Snitz, and Chad Nelson, "Clinical Versus Mechanical Prediction: A Meta-Analysis," *Psychological Assessment* 12 (2000): 19–30; John A. Swets, Robyn M. Dawes, and John Monahan, "Psychological Science Can Improve Diagnostic Decisions," *Psychological Science in the Public Interest* 1 (2000): 1–26.
- [19] Philip E. Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton University Press, 2005.)
- [20] Kesten C. Green, "Forecasting Decisions in Conflict Situations: A Comparison of Game Theory, Role-playing, and Unaided Judgement," *International Journal of Forecasting* 18 (2002): 321–44.
- [21] Two examples of this type of research are: Robert D. Folker, Jr., "Exploiting Structured Methodologies to Improve Qualitative Intelligence Analysis," unpublished masters thesis, Joint Military Intelligence College (1999); and Brant A. Cheikes, Mark J. Brown, Paul E. Lehner, and Leonard Adelman, "Confirmation Bias in Complex Analysis," MITRE Technical Report MTR 04B0000017 (2004).
- [22] The Columbia space shuttle investigation concluded: "The Board views the endemic use of PowerPoint briefing slides instead of technical papers as an illustration of the problematic methods of technical communication at NASA." Columbia Accident Investigation Board Report, Vol. 1 (August 2003), 191.
- [23] Sherman Kent, "Words of Estimative Probability," *Studies in Intelligence* 8 (1964): 49–65; David Budescu and Thomas Wallsten, "Processing Linguistic Probabilities: General Principles and Empirical Evidence," in Busemeyer, et al., eds., *Decision Making from a Cognitive Perspective* (New York: Academic Press, 1995).
- [24] For a different view of the analogy between intelligence analysis and medicine see Stephen Marrin and Jonathan Clemente, "Improving Intelligence Analysis by Looking to the Medical Profession," *International Journal of Intelligence and Counterintelligence*, 18 (2005): 707–29.
- [25] WMD Commission Report, 399.
- [26] *Ibid.*

Steven Rieber is a scholar-in-residence at the CIA's Kent Center for Analytic Tradecraft. Neil Thomason is a senior lecturer in history and philosophy of science at the University of Melbourne. This paper is an extract of a longer manuscript in process.