Naval War College Review

Volume 73
Number 1 *Winter 2020*

Article 8

2020

Research & Debate: Prediction

Wayne P. Hughes Jr.

RESEARCH & DEBATE

PREDICTION

Wayne P. Hughes Jr.

The predictive power of experts, operations analysis, and the value of information are interwoven subjects that are hard to winnow down to an essence. Prediction is a big subject, so I have limited this article to what I know best: the operational and tactical domains of conventional warfare.

First you will read three examples of limitations of predictions when they are formed on the basis of information alone. Next I will demonstrate that even a modest amount of quantitative analysis, even with incomplete information, can help a decision maker execute a military campaign without making explicit predictions about the coming battles or operations. Analysts cannot eliminate wartime surprises, but they can help to avoid the worst mistakes and steer military leaders toward better decisions. I will conclude by advocating what is too rarely done: the comparison of quantitative campaign analysis done before a war with what actually transpired in the war, to show that useful—even critically important—advice can be formulated very quickly to help decision makers. On one hand, intense thinking about the war is necessary; on the other hand, expert judgment alone should be augmented with simple, transparent, timely—even if incomplete—quantitative analysis.

Dean Emeritus Wayne P. Hughes Jr. served from 1982 to 2017 in many teaching and research capacities at the Naval Postgraduate School in Monterey, California. He also taught naval history and foreign policy at the U.S. Naval Academy, in Annapolis, Maryland, from 1957 to 1960. From 1952 to 1982 he was on active duty in the Navy, retiring at the rank of captain. He held several commands and served in six operations research billets, both afloat and in the Pentagon. Among his books are three editions of Fleet Tactics *and* On Naval Tactics *in USNI's "Wheel Book" series.*

Naval War College Review, Winter 2020, Vol. 73, No. 1

PREDICTION FROM INFORMATION ONLY

Black Swans

Surely the most drastic book on prediction is N. N. Taleb's *The Black Swan,* subtitled *The Impact of the Highly Improbable.* Taleb makes an entertaining

case for the occurrence of unforeseeable events and the like, but his advice is pretty trite; since, by definition, a black swan cannot be predicted, the most we can do is be ready for surprises, and then be responsive and adaptive when they occur.[1]

Gray Swans

More interesting are what might be called *gray swans:* surprising events of great consequence for which evidence existed beforehand but was lost in a clutter of information. In the commercial sector, the recent burst of the housing bubble is the latest of many collapses brought on by "the madness of crowds" whose herd instinct overcame many clues of excesses in plain sight.[2] Gray swans in the military domain are exemplified by the invasion of South Korea in 1950, the collapse of Soviet control in 1989, and the invasion of Kuwait by Iraq in 1990. All three illustrate "surprises" that Monday-morning quarterbacks have decried. After the debacle at Pearl Harbor was described alternatively as resulting from a nefarious plot or the careless handling of information, Roberta Wohlstetter wrote what is, to me, the definitive interpretation in *Pearl Harbor: Warning and Decision.*[3] At the strategic intelligence level, she shows that it was easy to miss the clues of what turned into a tactical disaster amid information overload. At the emotional level, one must see the need to hedge against human shortcomings in predicting future wars, while being careful not to cry wolf too often.

Gray swans are complicated by the fact that an enemy frequently will use deception to ensure they are gray, so to speak. In part Japan was successful in its Pearl Harbor attack because it employed deception to achieve surprise. I will refer later to Barton Whaley's masterful study of strategic deception, but here I will mention an equally valuable source: the recent book by Erik Dahl, *Intelligence and Surprise Attack: Failure and Success from Pearl Harbor to 9/11 and Beyond.*[4] Dahl is particularly insightful because he goes beyond deception in big wars to include deceptions that terrorists use to attempt to achieve surprise.

Expert Political Judgment

What, then, about predictions by experts? A marvelous book by Philip E. Tetlock first describes finding 284 self-proclaimed authorities who made a living commenting on political, international, or economic trends and were willing to participate in his study. Tetlock's questions were the kind that could be answered "better," "worse," or "about the same." Over several years in the 1990s Tetlock accumulated 82,361 answers in his database. In 2003, Tetlock compared the predictions with actual results. Two years later he published his conclusions in a book entitled *Expert Political Judgment.*[5]

And the envelope, please. Well, it is a fat envelope, because Tetlock gives all the interested parties a nuanced hearing. To summarize his findings, I quote from a *New Yorker* book review: "[t]he experts performed worse than they would have if

they had simply assigned an equal probability to all three outcomes. . . . Human beings who spend their lives studying the state of the world are poorer forecasters than a dart-throwing monkey."[6] Worse still, the experts tried to defend their wrong predictions with excuses such as "My timing was off" or "An unforeseeable event interfered with what should have happened." Tetlock also shows that nonexperts who answered the same questions did better than the dart-throwing monkey. Not a lot better—but significantly more so than the experts.

INFORMATION SUPPLEMENTED BY ANALYSIS

These are examples of predictability on the basis of information only. Tetlock's *Expert Political Judgment* is full of statistics measuring the performance of experts, but as far as I know the experts did not do any analysis to supplement their opinions. So, let us next make a distinction between information-based prediction alone and decision-making that is assisted by a quantitative assessment.

Why Military Analysis Cannot Predict

As we begin the shift to military operations analysis, I refer to an essay by the late, great Air Force analyst Clayton Thomas.[7] In effect, he described model-based analysis as an if-then statement. Two things—the model and its inputs—are on the if side; model and inputs together are processed to yield a result—the then side. If the model represented reality—which in campaign analysis it cannot—and if the data were precise—and in warfare the data are always "dirty" with errors—then the result would be an accurate prediction. We military analysts make no such claims; we say no more for the results than that when they are used wisely insightful conclusions can be reached and better decisions made.

Prediction Is Sometimes Unavoidable

Although generally we do not claim to predict outcomes, sometimes a decision maker must do just that, and we must help him. A prominent example is the procurement of warships and aircraft that are intended to have thirty- or even forty-year service lives. To see the impossibility of getting the designs right, no matter how comprehensive the analysis may be, reflect on the state of the world in 1979 and all that has changed since then that affects the prospective wartime performance of those ships today.

Space permitting, I could write at length, first, about how our warships completed before 1979 were designed earlier to carry technologies that were earlier still; second, that expensive, multipurpose ships are a poor way to hedge against future gray swans; and third, that we have not had to fight a fleet battle since 1945. All our learning about war at sea in the missile age has been vicarious, except for the handful of embarrassing single-ship attacks we have suffered.

Analytical methods and predictive power vary with tactics, technologies, and testing and with whether the predictions concern policy, operations, logistics, procurements, or strategies. A fine book on the subject is the Military Operations Research Society's *Military Modeling for Decision Making,* because it is comprehensive in distinguishing modeling and techniques for different defense-related purposes.[8]

Strategic Planning and Force Procurement

An accurate, recent, thirty-eight-page appraisal of predictive power when aided by extensive, even exhaustive, analysis was published in October 2011 by the distinguished statesman Richard Danzig.[9] Quoting liberally from both Taleb and Tetlock, Danzig shows the limits of model-assisted planning and why the limits have been inevitable when programming weapon systems for the future. His cure is difficult to implement, however, because Danzig argues, in black swan fashion, for more-nimble Department of Defense and congressional processes and acceptance of something less than the perfection demanded by those in government who metaphorically dodge and weave in a defensive crouch.

In one respect, Danzig's advice seems implementable. He recommends that we work on simpler systems that can be designed and produced more quickly and be discarded after shorter lifetimes, when geopolitical circumstances change or new technologies serve up either threats or opportunities. Although Danzig does not say it this way, the implication is that top-down solutions are unavoidable when expensive, long-lived systems must fill capability niches that will endure for the long haul—for example, multifunction orbiting satellites, or ballistic-missile-carrying submarines armed with "failure proof" nuclear weapons for strategic deterrence. Otherwise, bottom-up, quickly deliverable, relatively inexpensive systems that fill immediate needs—sometimes by short-circuiting the procurement bureaucracy—are the way to recover from failures of prediction in strategic planning. An example is the recent, rapid development of unmanned aerial vehicles, both in the large quantities deployed and in the many design variations. Falling somewhere in between were the successes at Kelly Johnson's Lockheed Skunk Works, which responded quickly—from the U-2 in 1955 to the SR-71 in 1966—to fill a need for long-range surveillance aircraft perceived at the highest levels of the Central Intelligence Agency.

Strategic Deception in Wartime

Barton Whaley's *Stratagem* is a good, quantitative book on methods of deception to achieve strategic surprise, how many false clues it takes to achieve it, how to enhance your chances of success, and why attempts to deceive have not cost much in resources.[10] He gives historical examples, such as the strategic surprise the

Germans achieved in 1941 when they invaded the Soviet Union and what the Allies achieved in the Normandy invasion. Whaley tells the deceiver how to succeed and the rewards that ensue. He shows that the victims of strategic deception behave much like Tetlock's experts, who were blinded by their own overconfidence.

Tactics, Technology, and Testing

The measured performance predictions in peacetime exercises become caught up in the fog of war. Jon Sumida observes that before World War I the Royal Navy expected a hitting rate of 30 percent with the fleet's big guns.[11] But in the Battle of Jutland the Germans achieved a rate of about 4 percent and the British 3.5 percent.[12] There were good reasons for the diminished performance at Jutland, but that is the point about prediction: there are *always* going to be ex post facto reasons your peacetime expectations will be wrong. The English operations analyst David Rowland has devoted much of his career to comparing ground combat exercise data with wartime data from similar battles. In one of his early papers he compares results from laser-instrumented, nonlethal training exercises with actual combat results in similar environments during World War II. The predictions based on the exercises overestimated the casualty production rate for tanks by a factor of two; for artillery duels by a factor of three; and for pure infantry actions by a factor of *seven*! Yes, a sevenfold overestimation of soldier performance. In effect, Rowland confirms S. L. A. Marshall's highly controversial conclusions about the small number of American soldiers who fired their weapons when under fire in World War II.

One of the most famous model-based predictions—I think *prediction* is the apt word—was by Frederick W. Lanchester, who claimed that the square law phenomenon would apply to air-to-air combat.[13] He wished to show the advantage of numbers over quality in a new age of air warfare. But Lanchester was wrong. From evidence reported by Philip Morse and George Kimball in their famous *Methods of Operations Research* and in more-detailed recent analysis by Niall MacKay, we know that through World War II the linear law applied in the air.[14] What Lanchester failed to see was that air combat essentially consists of duels, in the form of dogfights or ambushes, so the square law assumptions are not met. This was no theoretical matter. As MacKay shows, the top Royal Air Force leaders in the Battle of Britain argued between massing defending fighters—in Lanchester square law fashion—and getting the fighters in the air as swiftly as possible, so that small detachments were in the best position to win duels between single aircraft. I also tell our students of campaign analysis that the greatest number of kills often does not come from air-to-air combat. If they want to anticipate—to predict, as it were—which side will achieve air superiority, they must make a

difficult estimate of how successful each side's attempt will be to attack aircraft on the ground, the way the Japanese surprised and destroyed MacArthur's air force in the Philippines immediately after Pearl Harbor.

Lest you think we are better off now, with modern computers and powerful algorithms built into our best models, here is a more recent example. The U.S. Navy depends mightily for defense of the fleet on the Aegis missile system. Using data from controlled experiments at sea, one might conclude that if you shoot two surface-to-air missiles (SAMs) at an incoming antiship cruise missile (ASCM), and if you also add some point defense, you can expect to shoot down 90 percent or more of the attacking ASCMs. What is the combat record? In battles at sea, warships of other states have averaged around 75 percent success in defending themselves from ASCMs. But all the successes must be attributed to soft-kill and point-defense weapons, not to SAMs. There are also several instances of warships that might have defended themselves but did not, illustrated by the recent successful missile attack on the Israeli warship *Hanit*. Navy officers also will remember the Exocet hits on USS *Stark* and HMS *Sheffield,* which might have defended themselves with surface-to-air missiles but did not. In the entire record, starting in 1967, of more than 220 missiles fired on ships at sea, only one antiship missile has been shot down by a SAM. The record of USN missile ships in combat is zero for two, if one counts the action of USS *Vincennes* in shooting down an Iranian airliner as a failure. As at Jutland, a careful examination of these missile-era events shows there were reasons for the wartime results—pretty good reasons, too—but the important conclusion is that the fog of war almost always makes peacetime predictions too optimistic. Wartime surprises, although not exactly black swans, always will be present.

OUR PRODUCT IS USEFUL INSIGHT

Now I am going to focus on the domain of gray swans when our tools are used for operational and tactical predictions. I will show that even though the predictive power of our analyses is less than we would wish, if we focus on the right objectives and use appropriate measures of effectiveness our results and recommendations will be a powerful aid to decision makers. Indeed, I am going to arrive at conclusions so cheerful they may surprise you.

Campaign Analysis

Campaign analysis is hard to do, and its predictive power is very much a matter of how demanding you want to be. For example, between the world wars the Naval War College played over three hundred games, most at the campaign level and most against Japan. They were highly valuable because they sobered our early optimism about the war's most important elements. After the war, Admiral

Chester Nimitz wrote a famous letter saying that, except for kamikazes, the games had anticipated its major events accurately—referring, I suppose, to what happened in the drive through the Central Pacific that he oversaw. On the other hand, the Guadalcanal campaign, the shift from a battleship-centric force to a carrier-centric force, the vital contribution of American code breaking, and the drive by MacArthur up the New Guinea coast were vital aspects about which the games afforded no clues. In fact, after Pearl Harbor every class of warship except minesweepers changed its function.

At the tactical level, even the postmortems do not do justice to two factors that some operations analysis might have revealed. Looking back at the Battle of Midway of June 1942, historians recognized four things that were necessary for the Americans to overcome a numerical inferiority of seventy-five ships to twenty-five: code breaking; brilliant leadership by Nimitz, Fletcher, and Spruance; great courage in our naval aviators; and just plain good luck. But they missed two others. Until recently, no historian had picked up on the value of radar. If the Japanese ships had had our air-search radar, our surprise dive-bomber attack could not have succeeded.[15] Nor has any historian I have read identified the key role of Midway Island itself, which served as an immobile fourth American aircraft carrier, drawing away Admiral Chuichi Nagumo's attention and firepower at the critical time.[16]

And yet, and yet: war games and fleet exercises schooled our carrier commanders before the Pacific War to know that the best way to win—and the only way, if outnumbered—was to detect the enemy first and get off a decisive first strike with every aircraft you had. Simple but elegant salvo equations, not yet invented in 1942, would match the results and "predict" with sufficient quantitative accuracy the outcomes of all five of the big carrier battles in the Pacific ex post facto.[17]

Having in mind, then, that both Admiral Nimitz and the Midway historians ought to be given some slack, I will now describe three remarkable examples of the power and utility of our methods applied to campaigns, to show how analyses can help military leaders make better decisions and avoid the worst blunders. The examples are entertaining because they were performed by our young officer students at the Naval Postgraduate School (NPS) in a course on joint campaign analysis. The students had to reach their conclusions very quickly, with maximum professional knowledge and minimum computation, because the class pretense—a realistic one—was that their decision maker needed their inputs within about seventy-two hours. In these "ministudies," the students did not have time to construct a detailed, realistic simulation.

Foresight and Hindsight in Wars

The Falkland Islands War. In the first example, the students fought the Falklands War on paper in 1982 before it started. They had no inkling that *General Belgrano* was about to be sunk, taking the Argentine surface navy out of the war; or that Exocet missiles would be highly effective in destroying British ships; or that Argentine ground forces in the Falklands would be thoroughly outclassed. To do justice to their insightful work would take several paragraphs, but I can report the bottom line very quickly.

Neither side that fought had done such an analysis—early, fast, and basic. I believed then, and still do, that if they had the Argentine junta would have won the war, and British prime minister Margaret Thatcher would have been more cautious about sailing forty-two ships—essentially committing the United Kingdom to take back the Falklands. Why? Because the focused campaign analysis by the students showed that, with only a little foresight, the Argentine air force—all 140 fighter/attack aircraft, flown by capable pilots—could have staged through Stanley airfield on East Falkland Island. It did not take a detailed model or precise inputs to conclude that those aircraft, even when dropping iron bombs the old-fashioned way, would have penetrated the twenty-two Harriers and other air defenses and put enough British ships out of action to force the fleet to abandon the invasion.

Operation Desert Shield. While my class was meeting in the fall of 1990, a big debate was raging over whether the United States and our Middle East partners could force Saddam Hussein out of Kuwait without a ground campaign. At the time, many American members of Congress and pundits were arguing that this was possible.

The charge to my students was to do a fast-turnaround ministudy to determine whether there was an operation other than an invasion of Kuwait that would persuade Saddam Hussein to leave. After doing as much quantitative assessment as time permitted, the students concluded that if we wanted him out of Kuwait we would have to attack on the ground. This seems obvious in hindsight now, but it was not so when the students made their appraisal.

Operation Iraqi Freedom. Lastly, I report on Operation Iraqi Freedom (OIF). This student appraisal was done even faster than a ministudy—over a single weekend. We asked the students how long it would take to win the war. Astutely they asked, "What do you mean by 'win the war'?" Together we agreed that getting to Baghdad and toppling Saddam Hussein would constitute victory! I still think that was a suitably specific analysis goal, because everything after that comprised peacemaking operations—long and difficult though they turned out to be.

Four student teams each made independent estimates. One team said it would take four weeks; one team said two to four weeks; one team said two weeks to get there, but that they did not know how long the city fighting would last; and the last team said three weeks. As it transpired, it took three weeks and a day to overthrow the regime. What our students could not predict, of course, was that a sandstorm would slow the advance, and that some elite soldiers operating inside Baghdad would enjoy such luck and display such courage.

But the students had some crib notes to help them make their estimates. They knew that research, most notably by the Army analyst Bob Helmbold, had concluded that the rate of advance of an army unopposed or against light opposition has been and still is about twenty-five miles a day. The students could scale back the movement rate appropriately in making their estimates—predictions, as it were. In actuality, our soldiers and Marines advanced the three hundred miles to Baghdad in three weeks—a rate of fifteen miles per day.

OTHER DOMAINS

I have emphasized the rewards and limitations of operational and tactical analysis to prepare for war. There is a lot more to the story. Before summing up, here is a brief contrast with two other domains of prediction.

Attenuating Terrorist Attacks

I am not well informed on what kind of analysis would best supplement experience in fighting the perpetual war against terrorists. But I have read a fine paper entitled "How Probabilistic Risk Assessment Can Mislead Terrorism Risk Analysts." It is a warning against a methodology that cannot help and might hinder prediction and planning for homeland defense.[18]

Authors Jerry Brown and Tony Cox see two problems with the methodology. One danger is to put confidence in the predictions of experts that are in fact inputs to the analytical scheme. They are suspicious of expert opinion, as am I. The other problem is adapting a methodology—probabilistic risk assessment (PRA)—that has been effective for engineering analysis but is essentially a decision-theory way to design against adverse natural events and risks. The authors point out that when the enemy is not nature but an attacker who wants to outwit us and penetrate our defenses, then PRA actually can help the enemy. The proper mind-set is game theory, which says that we must do the best we can against the best he can do. The PRA methodology comes no closer to examining enemy choices and capabilities than to ask an expert the "probability of an attack"—without regard for what the enemy observes us doing.

There is wide agreement that a new attack against our homeland will come someday. Predicting where and against what target is the hard part that the PRA

method cannot illuminate. Brown and Cox recommend shifting the emphasis of risk management from using experts to guess where risk might be greatest to calculating where targeted investments will most improve the resilience of critical infrastructures. This entails more attention to two things: First, install additional safeguards where they have the biggest payoff—for example, by adding some "inefficient" redundancies to our excessively "efficient" but vulnerable electrical-distribution system, as well as the grids that distribute trains, trucks, petroleum, and communications. Second, establish and practice procedures to recover after an attack—for example, on the large containerport at Long Beach or the San Francisco–Oakland Bay Bridge. Perhaps we have improved disaster recovery since the terrorist attack on the Twin Towers, but from the natural disasters of which I am aware, such as the 2004 Indonesian tsunami and Hurricane Katrina, and also from an NPS-conducted experiment in intergovernmental cooperation in San Diego, preparing to act after an attack may be more productive than trying to prevent every attack. It is likely that the two best ways to recover more quickly are by conducting inexpensive drills to improve coordination among many agencies and levels of government, and by ensuring readiness to employ emergency modes of communication. The general rule is "when there's a war on, study the war." That applies to the war on terrorists, to the frequent use of unmanned vehicles in peacetime, and to the unending competition to safeguard and exploit cyberspace.

Measuring Influence to Avoid War

The object of the Cold War was to exert American influence without fighting the Soviet Union. We never could measure past success in predicting outcomes of our campaigns, including a highly predictable world disaster from a nuclear exchange, because there was no war to study. The paradox is that the only available measures of the success of the analyses and the predictions resulting therefrom were that, year by year, deterrence held.

As far as I know, during the long Cold War there was only one attempt to measure the predictive power of the many campaign analyses of a hot war. It occurred because an inspired analyst at the Center for Naval Analyses (CNA) persuaded the CNA president to refight, on paper, a study that his think tank had conducted for the Navy in 1965, of a war at sea conducted ten years later. A study assumption was that the nuclear threshold would not be breached, in part because the American strategy was to confine the war to the oceans. The war was bloody enough among the combatants, but massive civilian casualties were avoided. Around 1976 (I am citing from memory) the analysis was repeated with the same military objectives, but with the geopolitical environment updated, resulting in somewhat heavier demands on NATO forces, principally those of the

U.S. Navy. In 1976, Soviet forces were slightly greater than had been projected in 1965; our estimates of Soviet maritime combat capabilities were about the same. On the other hand, between 1965 and 1976 the American fleet had become much smaller, and future combat capabilities projected to be in the fleet in 1976 had not lived up to technical expectations when actually deployed. Thus, all inputs for the repeat campaign analysis seemed to indicate a worse outcome. Yet the outcome of the campaign "fought" on paper in 1976 was amazingly about the same as in the 1965 study, and perhaps a little bit better. The reason was that in two or three instances after the new systems built in 1965 were deployed, new tactics were conceived and developed to fight with them more effectively. Better tactics more than offset technological disappointments and our smaller fleet.

But that interesting finding is peripheral to the two main points. First, it is highly useful to test our tactical and campaign analyses when their inputs and assumptions can be tested, yet it hardly ever is done. Second, the study results—even in 1965, and despite their flaws—were decisively instructive. The purpose of the study was to test whether a war against the Soviet navy limited to the sea was attractive for NATO. The answer was no. That was the conclusion of overarching importance. As with the Falkland Islands scenario, it did not take exquisitely detailed analysis to see why—after the analysis had been done. The Soviet Union was a continental power that did not depend fundamentally on the oceans, but NATO was a maritime alliance for which control of the Atlantic was essential. The Soviets had too little at risk and NATO too much risk to make the threat of a war at sea an effective deterrent. No more was heard of it. NATO continued, wisely, to believe the central front in Europe was the critical region of interest.

There is a modern analogy to the war at sea, as we contemplate ways to influence China, keep faith with friendly states in Asia, and avoid a big and economically disastrous war. Far from being unwise, analysis of a war-at-sea strategy in the western Pacific looks feasible and desirable because, unlike that of the Soviet Union, Chinese prosperity depends on the sea. Unlike the Soviet state, China has begun to build a fleet that can protect the movement of its shipping in the open ocean, while shifting from a sea-denial to a sea-control navy.

A brilliant recent article by Naval War College professors Toshi Yoshihara and James Holmes points out that one cannot construct a strategy unless its *ends, ways,* and *means* are well defined. Hence the state—namely, China—must be identified as the strategy's object.[19] The *ends* almost have been established, because the Secretary of State and others have indicated our intention to put more emphasis on the western Pacific. In effect, our policy experts have made a prediction about the future. Next must come an analysis of the best *ways* to sustain our influence there at an affordable cost. Yoshihara and Holmes describe the limits

of the Air-Sea Battle concept and suggest other actions at sea that can and should precede strikes on mainland China. U.S. and allied navy components should try to keep the war at sea, exploit American maritime strengths, and demonstrate that China has the most to lose at every level of escalation—from maritime interdiction short of a full blockade all the way up to sinking Chinese warships and commercial vessels with American submarines in their home seas. Once we have in hand the ways to constrain every kind of confrontation, next comes further campaign analysis, testing, and negotiation with allies and partners in Asia. We must ascertain the *means:* the types and numbers of forces to execute such a flexible strategy that also fits the desires of China's neighbors and worldwide commercial interests.

The same fleet must be suitable in times of cooperation, competition, confrontation, or conflict, and China has a say in what our ends must be. If all our ships are expected to have thirty- and forty-year service lives, the challenge will be to construct one long-lived fleet for all circumstances. We do not yet know whether Yoshihara and Holmes are right about the ways and means, but analysis to meet various conditions, not a prediction of a single future, is the way to find out.

WHAT TO BELIEVE ABOUT PREDICTION

Black swans exist. Unavoidable surprises will continue. Black swans do not have to be deceptive, because, by definition, their surprise cannot be predicted.

Gray swans in the military world are complicated because they are concealed by a perverse enemy who wants to surprise us. Pearl Harbor happened not just because it was an unlikely event and the clues about the attack were mishandled but also because a clever enemy was doing his utmost to surprise us.

Regrettably, gray swans are not likely to become rarer. The growth of knowledge, illustrated by the replacement of a written *Encyclopaedia Britannica* with the electronic Wikipedia, exceeds our capacity to sort the information quickly. And in fast-moving military operations the enemy will be trying constantly to throw sand in our eyes.

Expert judgment for national policy and military strategy is unreliable unless it is substantiated with the quantitative methods of operations research. Critical decisions can be greatly—even decisively—enhanced by quantitative analysis, notwithstanding that a decision maker's prewar conclusions will fall well short of—and should never claim to be—a prediction of the future. Useful insights come from wise application of dirty data processed in an appropriately simple model to yield results that are at once precise, inaccurate, and helpful.

The Falklands War seventy-two-hour analysis by our campaign analysis students illustrates how decisive macro insights can be discerned in a very short

time. Despite limited knowledge of how a war will unfold, quantitative analysis is powerful for uncovering the essential features of good and bad choices.

The students' overnight analysis ahead of OIF showed two different things. On one hand, the analysis demonstrated that an amazingly accurate estimate of how long it would take to overthrow the Hussein regime could be made swiftly. On the other hand, analytical power did not help anticipate that after toppling the regime there would be a very long aftermath of difficult peacemaking. It is not new news that the enemy gets a vote, and sometimes his choice will seem not to be in his own best interests.

Accurate predictions are useless if they are too late to help the decision maker. If he or she must act in seventy-two hours, we must help him within seventy-two hours. Our students follow the one-third, one-third, one-third rule of analysis. Given three days to complete the work, spend the first day figuring out how quantitative analysis can best help him make his decision; do the analysis on the next day; and take the third day to recover from your mistakes, answer his questions, or enrich the work.

There are many variations of conflict in which military operations analysis can supplement professional knowledge profitably. One is when the campaign goes on endlessly; this allows time to gather "combat" data, assess it, and apply it—while remembering that the enemy also is observing and adapting. Another is when the object is not to prepare for war but to adapt new ends, ways, and means to prevent war by retaining influence over a prospective enemy in changing circumstances. Then the goal of analysis is to help decide what strategy and capabilities will be the best ones to keep the peace or to contain the war at a low level of violence.

A paradigm of all prediction is the if-then statement, with two parts to the if side. To the extent that a model describes the circumstances and the data are accurate, the analysis process will give accurate results. When the model is a simplification (an artful one, we hope) and the data are dirty (but good enough, we hope), then the goal is not to predict the outcome but to help a decision maker do the best he can after adding his own wisdom to our quantitative analysis.

NOTES

1. Nassim Nicholas Taleb, *The Black Swan: The Impact of the Highly Improbable* (New York: Penguin Books, 2007).
2. I am referring to the classic book on economic bubbles by Charles Mackay, *Extraordinary Popular Delusions and the Madness of Crowds* (London: Richard Bentley, 1841).
3. Roberta Wohlstetter, *Pearl Harbor: Warning and Decision* (Stanford, CA: Stanford Univ. Press, 1962).

4. Erik Dahl, *Intelligence and Surprise Attack: Failure and Success from Pearl Harbor to 9/11 and Beyond* (Washington, DC: Georgetown Univ. Press, 2013).
5. Philip E. Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton Univ. Press, 2005).
6. Louis Menand, "Everybody's an Expert," *New Yorker,* 5 December 2005, pp. 98–100.
7. Clayton Thomas, "Verification Revisited," in *Military Modeling for Decision Making,* ed. Wayne P. Hughes Jr. [Capt., USN (Ret.)] (Alexandria, VA: Military Operations Research Society, 1997).
8. Hughes, *Military Modeling for Decision Making*.
9. Richard Danzig, *Driving in the Dark: Ten Propositions about Prediction and National Security* (Washington, DC: Center for a New American Security, 26 October 2011), available at www.cnas.org/.
10. Barton Whaley, *Stratagem: Surprise in War* (1969; repr. Norwood, MA: Artech House, 2007). A valuable companion, also rich in quantitative and qualitative detail, is Donald Daniel and Katherine Herbig, *Strategic Military Deception* (New York: Pergamon, 1982).
11. Jon T. Sumida, "A Matter of Timing: The Royal Navy and the Tactics of Decisive Battle, 1912–1916," *Journal of Military History* 67, no. 1 (January 2003), pp. 106–107.
12. John Campbell, *Jutland: An Analysis of the Fighting* (Annapolis, MD: Naval Institute Press, 1986), pp. 346–55.
13. Frederick W. Lanchester introduced his theory in 1916 in *Aircraft in Warfare.* A more respectable and durable essay on the power of concentration treated more broadly is his "Mathematics in Warfare." See James R. Newman, *The World of Mathematics* (New York: Simon & Schuster, 1956), pp. 2138–57.
14. For example, see Niall MacKay, "Is Air Combat Lanchestrian?," *Phalanx* 44, no. 4 (December 2011). My Navy heritage demands that I say that the square law was appropriate for battles at sea in the battleship era because square law conditions were met. Also, two USN officers, J. V. Chase and Bradley Fiske, invented the square law to describe the advantage of numbers quantitatively, and they did so a decade before Lanchester, who never saw their work.
15. The Japanese had forty-one fighters in the air to protect the carriers, but they were at low altitude after shooting down the American torpedo bombers that served as an inadvertent but critical decoy.
16. I believe the journalist Hector Bywater's book *The Great Pacific War, 1931–33,* written in 1926, was nearly as good in predicting how the war would transpire as were all the Naval War College (NWC) games. Bywater's descriptions of the battle in a Pacific campaign had prescient predictions that even our many NWC war games did not foresee.
17. See Wayne P. Hughes Jr. [Capt., USN (Ret.)], *Fleet Tactics: Theory and Practice* (Annapolis, MD: Naval Institute Press, 1986), pp. 93–103.
18. Gerald G. Brown and Louis Anthony Cox Jr., "How Probabilistic Risk Assessment Can Mislead Terrorism Risk Analysts," *Risk Analysis* 31, no. 2 (January 13, 2011), pp. 196–204.
19. Toshi Yoshihara and James R. Holmes, "Asymmetric Warfare, American Style," U.S. Naval Institute *Proceedings* 138/4/1,310 (April 2012).