

THE LOGIC OF INTELLIGENCE FAILURE

Bruce G. Blair
President
Center for Defense Information
www.cdi.org
April 2004

INTRODUCTION

The severe post 9/11 criticism of the U.S. intelligence system for underestimating the terrorist threat to America, and for overestimating weapons of mass destruction in Iraq, would be sharply tempered if critics understood the laws and limits of reasoning. Uncertain threats tend to be misestimated initially, and only repeated assessments can close the gap between threat perception and reality. Even when the strict rules of inductive reasoning are applied to spy data, ten or twenty successive reviews are typically needed to ensure that perceptions match reality.

Critics presume that far fewer assessments should suffice, and accuse users of intelligence with dogmatism if they do not respond with alacrity to the first alarm bells warning of a rising threat, or to the latest report discounting a threat. This criticism implies that intelligence analysts should suspend their prior beliefs and seize upon only the latest intelligence inputs. At the same time, if the inputs prove to be wrong, critics blame intelligence analysts for not seeing beyond the evidence and divining intentions.

While intelligence analysts cannot be psychics, psychology does, and should, figure prominently in the process of interpreting intelligence. Subjective opinion and preexisting beliefs, held by intelligence analysts and users of finished intelligence, including the top national security decisionmakers, are core elements of reasoned interpretation. The key to success or failure in interpreting intelligence information lies in rationally adjusting prior beliefs to make them conform to incoming intelligence information.

Prior opinion plays a critical role in every intelligence endeavor associated with current national security priorities: avoiding accidental nuclear war, detecting weapons of mass destruction, anticipating terrorist attacks, and preempting America's enemies. The initial bias of decisionmakers can be a blessing or a curse, but all that we can reasonably expect is that it is properly revised as new intelligence arrives.

An argument can be made that the processing of intelligence followed laws of reason in the cases of 9/11 and Iraq's weapons of mass destruction. Applying a rule of logic known as Bayes' law to these cases shows that the intelligence process produced conclusions that were not only plausible but reasonable.

AVOIDING ACCIDENTAL NUCLEAR WAR

An illustration of the dramatic effect of initial opinion on intelligence interpretation is a hypothetical situation in which top leaders with their fingers on the nuclear button receive indications of an incoming nuclear missile attack.

The most dangerous legacy of the Cold War is the continuing practice of Russia and the United States of keeping thousands of nuclear weapons on high alert, poised for immediate launch on warning. The danger is that false indications of an incoming enemy missile strike could produce a mistaken launch in “retaliation.”

The need to react rapidly under the time pressures of incoming submarine missiles with flight times as short as 12 minutes and land-based missiles capable of flying half way around the globe in 30 minutes would be strongly felt from the top to the bottom of the U.S. or Russian nuclear chain of command. In order to unleash retaliatory forces before they and their command system are decimated by the incoming missiles, the early warning sensors (satellite infra-red and ground-based radar sensors) must detect the inbound missiles within seconds after their firing, and the detection reports must be evaluated within several minutes after they are received. That is the current requirement for the warning crews stationed deep inside Cheyenne Mountain, Colo. Then the president and his top nuclear advisors would convene an emergency telephone conference to hear urgent briefings from the warning team and from the duty commander of the war room at Strategic Command, Omaha, which directs all U.S. sea-, land-, and air-based strategic nuclear forces. The Stratcom briefing of the president’s retaliatory options and their consequences has to be accomplished in a mere 30 seconds (a longstanding procedural requirement), and then the president would have between zero and 12 minutes to choose one. A launch order authorizing the execution of this option would flow immediately to the firing crews in underground launch centers, in submarines, and in bombers, and within three minutes, thousands of nuclear warheads would be lofted out of silos toward their wartime targets, followed ten minutes later by many hundreds of nuclear warheads atop submarine missiles ejected from their underwater tubes.

These pressure-packed timelines reduce decisionmaking to checklists, and increase both the likelihood and the consequences of human and technical error in the nuclear attack warning and command system. Ironically, the risk of false warning of an incoming missile attack has actually been increasing since the end of the Cold War as a result of the steady deterioration of the Russian early warning network. Both its satellite and ground-based sensors have fallen into disrepair, and the human organizations that operate the network have been weakened by economic and social stresses and inadequate training.

There is an offsetting factor of crucial significance, however. While the risk of false warning has increased, the danger that Russia or the United States would actually launch on that false warning has declined dramatically. The reason is that the leaders of these

two countries would presumably heavily discount if not entirely dismiss reports of an attack, simply because the reports would be so incredible.

Russia and the United States are no longer enemies. That either country would deliberately attack the other is so utterly implausible that a neutral observer would rightly suppose that their top leaders would rise above the noise, emotion and time pressure of a reported incoming nuclear strike. These leaders cannot mechanically tie their actions to any warning and intelligence network, however highly touted it may be. At their lofty pay grade, what they think of the warning information would be inevitably and properly weighed by the background information they bring to it. Their prior opinion about the other side's good or ill intentions must be brought to bear on the situation, and that prior opinion today surely would cause them to disbelieve the warning and delay the fateful decision long enough to discover that the alarm was indeed false. On the other hand, a continuing stream of attack indications from multiple reliable warning sensors would compel a rationally calculating leader to believe that in all likelihood an attack actually is underway. The stream of data would compel a dramatic revision of the initial disbelief until the harsh reality sank in.

In other words, the effect of prior beliefs and psychology on the process of nuclear decisionmaking is very great in the context of launching nuclear missiles on warning that an attack is underway with missiles in the air. That was true during the Cold War, and it is true today.

PREEMPTING (PREVENTING) ENEMY ATTACK

The psychology of decisionmaking is even more pivotal in a context of launching counterattacks before any opposing missiles have been fired. Anticipating a first strike by a nation or group before the strike has actually started involves a certain amount of conjecture and demands a more careful screening of more ambiguous intelligence. Human factors are thus especially important today in the context of counterproliferation and homeland defense under the new national security strategy of the United States announced in September 2002 by the Bush administration.

This new strategy elevates preemption from the level of tactics to the level of strategy. It assumes that rogue states and terrorist groups cannot be reliably deterred, and therefore must be neutralized before they pose a clear threat of imminent attack. The strategy seeks to prevent America's enemies from acquiring weapons of mass destruction in the first instance, using U.S. military force if necessary, and seeks to disarm them after they have acquired such weapons, whether or not their use against the United States is imminent.

Because this strategy seeks to eliminate incipient threats before they materialize full blown, preemption is a misnomer, a mischaracterization. The strategy embraces preventive war as much as preemptive attack. It even covers the case in which the U.S. would attack a putative adversary before the adversary realizes it is going to attack the United

States – a wag would say that the idea in this case is that the United States would help the adversary make up its mind about attacking the United States by attacking the adversary first.

The new U.S. strategy is actually not so new. It is reminiscent of U.S. nuclear thinking in the early days of the Cold War when the United States was trying to figure out how to deal with the original “rogue” state developing weapons of mass destruction – the Soviet Union. President Bush’s new strategy is a throw back to the 1950s and 1960s when the United States was not yet prepared to accept deterrence as the primary, let alone sole, basis of U.S. security vis-à-vis the Soviet Union. The United States security establishment considered and pursued every option under the sun in addition to deterrence – preemption, preventive war, surgical decapitation strikes, counterforce first strike, missile defense, bomber defense, civil defense (homeland defense), and even covert special operations to assassinate key leaders.

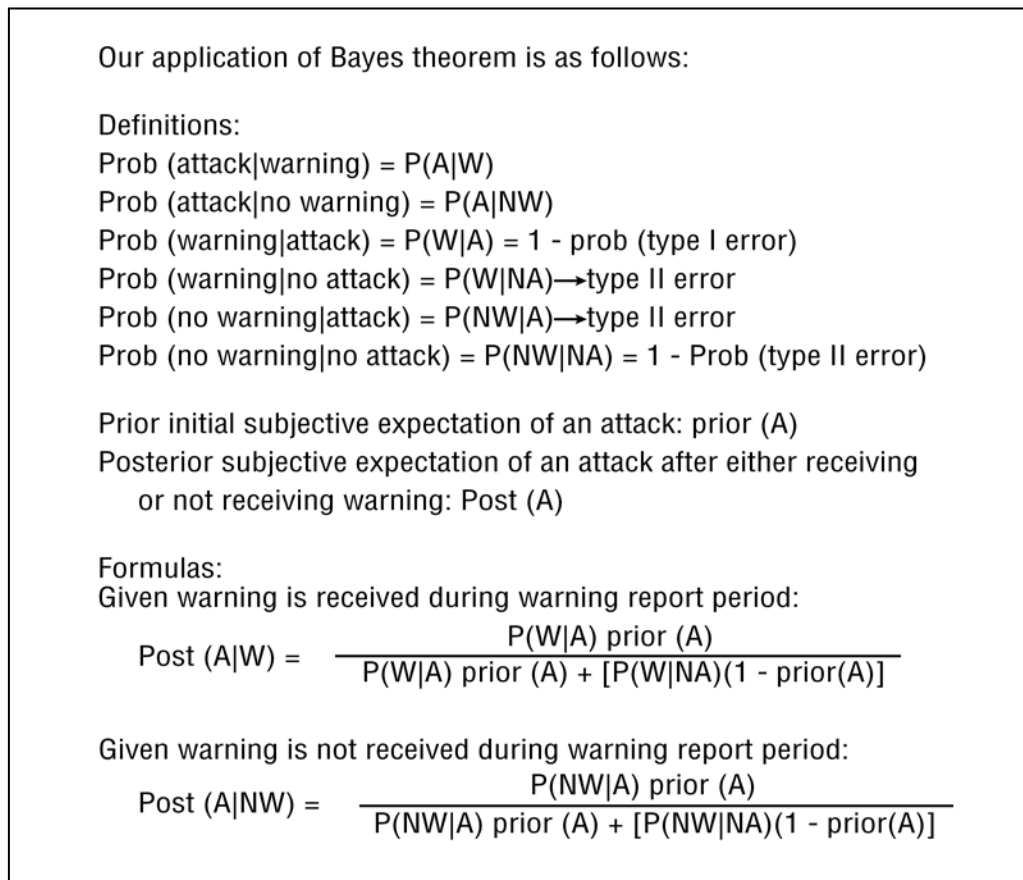
In the end, the U.S. and Russian security establishments realized that they could not meaningfully protect their countries and citizens from devastating strikes by the other side. None of the multitude of options being pursued could prevent either side from destroying the other in a nuclear war. Mutual vulnerability, despite intermittent attempts to remove it through Star Wars defenses or some other scheme, was a constant of the Cold War confrontation. But instead of despairing, both countries discovered salvation in this predicament. They were forced to rationalize mutual vulnerability as a virtue and learn to live with mutual deterrence as the centerpiece of national security, and eventually they celebrated this newfound source of security.

In contrast to this Cold War experience, however, the U.S. security establishment so far has rejected out of hand the idea of basing U.S. security on deterrence alone in confronting the far weaker axis of evil countries and terrorists. For understandable reasons, the United States is pursuing the same old options to protect itself from the rogue threats – active and passive defense and offense in line with the mindset of the early Cold War period.

A list of criticisms of the current U.S. preemptive strategy could run for pages. Its defects range from its dubious legitimacy under international law, to the bad example it sets for other countries eager to justify a preemptive or preventive attack on their neighbors. Already we have seen Russia and France follow in America’s footsteps to declare similar doctrine for themselves, and the list of emulators will undoubtedly grow.

High on this list of liabilities is one particular difficulty that is the focus of this essay: the enormous burden that preemption places on intelligence – not only intelligence collection and analysis, but its interpretation by those at the top who, as noted earlier, inevitably filter the intelligence information they receive through their own presumptions. The buck stops at a level at which leaders must fuse incoming intelligence with their own prior beliefs. It is crucial to the shaping of U.S. security policy that this highly subjective process be understood well. Intuition suggests that human intellectual and psychological limitations undercut the feasibility and sensibility of a preemptive strategy.

What is needed is a rigorous approach to analyzing whether the top leaders can interpret intelligence with sufficient accuracy and speed to meet the demands of the new strategy, even assuming that high-quality intelligence information can be collected and analyzed at lower levels. One such rigorous approach is to apply a proven formula for estimating the probability of an event – Bayes’ formula for contingent probabilities. This formula (see Figure 1) provides an account of how the required judgment, or interpretation, might be made in a disciplined, responsible manner. Bayes’ formula shows how well a perfectly rational individual can perform, providing a measure of the best judgment that can be expected of leaders in interpreting intelligence.



[Figure 1]

Bayes’ analysis is often called the science of changing one’s mind. The mental process begins with an initial estimate – a preexisting belief – of the probability that, say, an adversary possesses weapons of mass destruction, or that an attack by those weapons is underway. This initial subjective expectation is then exposed to confirming or contradictory intelligence or warning reports, and is revised using Bayes’ formula. Positive findings strengthen the decisionmaker’s belief that weapons of mass destruction exist or that an attack is underway; negative findings obviously weaken it. The degree to which the initial belief is increased or decreased depends on the intelligence system’s assumed rate of error – its rate of detection failure and its rate of false alarms. Bayes’ formula takes

both rates of error – known as type I and type II – into account in re-calculating probabilities.

All prior and posterior probabilities are strictly subjective in the Bayesian model. They are opinions that exist in the minds of individuals. Assessments supplied by intelligence and warning sensors do not objectively validate the probabilities, but merely enable existing opinion to be revised logically by the successive application of Bayes' formula. This process can be considered objective, however, in the sense that as more intelligence assessments based on real data become available, the subjective probabilities will eventually converge on reality. People with different initial beliefs will eventually agree with each other completely, if they are thinking logically. This consensus will be reached faster if the intelligence system is not prone to high rates of error.

Two Hypothetical Cases: Iraq's WMD and 9/11 Terrorist Threat

How subjective probabilities should be revised logically, according to Bayes' formula, are illustrated below for two hypothetical cases. One case resembles the problem of overestimating Iraq's weapons of mass destruction, and the other resembles the pre-9/11 intelligence failure in which a terrorist threat was underestimated.

In the case akin to pre-war Iraq, suppose that the national leader believes that dictator X is secretly amassing nuclear, biological or chemical weapons, but that U.S. spies cannot deliver the evidence proving the weapons' existence. What should the leader believe then? Should the indictment be thrown out if the spies cannot produce any smoking guns? How long would a reasonable person cling to the presumption of the dictator's guilt in the absence of damning evidence?

The mathematics of rationality (according to Bayes) throws surprising light on this question. It proves that a leader who continues to strongly believe in the dictator's guilt is not being dogmatic. On the contrary, it would be irrational to drop the charges quickly on grounds of insufficient evidence. A rational person would not mentally exonerate the dictator until mounting evidence based on multiple intelligence assessments pointed to his innocence.

The extent to which a rational person should change their mind about guilt and innocence depends on how reliably accurate the intelligence system normally is. Let's suppose the track record of the system suggests that it normally detects clandestine proliferation in 75 percent of the cases, and also that it avoids making false accusations in 75 percent of the cases. Thus, it misses proliferation in one-fourth of the cases, and mistakenly cries wolf in one-fourth of the cases. These rates of error seem to be reasonable approximations of current U.S. intelligence performance in monitoring clandestine proliferation.

If the leader interpreting the intelligence reports holds the initial opinion that it is virtually certain that the dictator is amassing mass-destruction weapons – an opinion that may be expressed as a subjective expectation or probability of, say, 99.9 percent – then what

new opinion should the leader reach if the intelligence community (or the head of a UN inspection team) weighs in with a new comprehensive assessment that finds no reliable evidence of actual production or stockpiling?

Adhering to the tenets of Bayes' formula, the leader would combine the intelligence report with the previous opinion to produce a revised expectation. Upon applying the relevant rule of inductive reasoning, which takes into account the 25 percent error rates, the leader's personal subjective probability estimate (the previous opinion) would logically decline from 99.9 percent to 99.7 percent! (see Figure 2). The leader would remain highly suspicious, to put it mildly, indeed very convinced of the dictator's deceit.

Initial and Revised Expectations of Hidden WMD (Given No Detection) Assuming a Detection System with 25 Percent Types I and II Error Rates.

Initial estimate ^a	Revised estimate given no detection													
	Number of Negative reports:													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0.0001	0.000													
0.001	0.000													
0.01	0.003	0.001	0.000											
0.05	0.017	0.006	0.002	0.001	0.000									
0.10	0.036	0.012	0.004	0.001	0.000									
0.20	0.077	0.027	0.009	0.003	0.001	0.000								
0.30	0.125	0.045	0.016	0.005	0.002	0.001								
0.40	0.182	0.069	0.024	0.008	0.003	0.001								
0.50	0.250	0.100	0.036	0.012	0.004	0.001								
0.60	0.333	0.143	0.053	0.018	0.006	0.002	0.001							
0.70	0.438	0.206	0.080	0.028	0.010	0.003	0.001							
0.80	0.571	0.308	0.129	0.047	0.016	0.005	0.002	0.001						
0.90	0.750	0.500	0.250	0.100	0.036	0.012	0.004	0.001						
0.95	0.864	0.679	0.413	0.190	0.073	0.025	0.009	0.003	0.001					
0.99	0.971	0.917	0.786	0.550	0.289	0.120	0.043	0.015	0.005	0.002	0.001			
0.999	0.997	0.991	0.974	0.925	0.804	0.578	0.314	0.132	0.048	0.017	0.006	0.002	0.001	
0.9999	0.999	0.997	0.992	0.976	0.932	0.821	0.604	0.337	0.145	0.053	0.016	0.006	0.002	0.001

a. Degree of belief in the hypothesis "weapons of mass destruction exist."

[Figure 2]

A leader believing so strongly in the correctness of that judgment might well order another independent intelligence review, expecting that it would produce positive findings this time around. Suppose that this review, much to the leader's surprise, repeats the earlier negative findings - no reliable evidence of weapons proliferation. What new opinion should the leader form then? A rationally calculating person would undergo another change of opinion after absorbing the second intelligence report, revising downward again, this time dropping from 99.7 percent to 99.1 percent. Believe it or not, a rational leader could receive four negative reviews in a row from the spy agencies and would still harbor deep suspicion of the dictator because the leader's logically revised degree of belief that the dictator was amassing weapons would only fall to 92.5 percent.

This seemingly dogmatic view is in fact the logically correct one. Why? Because top leaders do not function in a contextual vacuum. They inevitably depend on their own presumptions. And in the Iraq case, a very strong initial presumption of guilt is understandable in view of the regime's history. In late 1998, UNSCOM issued its final report listing WMD capabilities that remained unaccounted. Iraq still had not disclosed those capabilities fully in its December 2002 report to the United Nations. In view of this failure and of Iraq's historical intentions to acquire WMD, it's not surprising that leading up to the U.S. invasion of Iraq in 2003 the overwhelming bipartisan expert consensus of the United States and practically all other nations with modern intelligence capabilities was that Iraq certainly possessed at least a stockpile of chemical and biological agents.

Nobody seriously challenged that assessment, and if the rational calculations discussed above bear any resemblance to actual intelligence assessment during this period and after the war, it is no surprise that many of the most informed experts to this day still cling to the belief that Iraq possesses such weapons. Exhibit "A" is the recent public defense of the infamous National Intelligence Estimate of October 2002 mounted by the key CIA official responsible for its conclusion that Iraq had chemical and biological weapons. As Stuart Cohen, the official in question, puts it in his closing editorial comment.

"Men and women from across the intelligence community continue to focus on this issue because finding and securing weapons and the know-how that supported Iraq's WMD programs before they fall into the wrong hands is vital to our national security. If we eventually are proved wrong – that is, that there were no weapons of mass destruction and the WMD programs were dormant or abandoned – the American people will be told the truth; we would have it no other way."

(The Washington Post, "Myths About Intelligence," Nov. 28, 2003, P.A41).

In the case of the Sept. 11 attacks, the initial apprehension of suicide attack using hijacked planes against buildings was as low as the Iraqi WMD threat estimate was initially high. The terrorist strikes came as such a total surprise that the furious criticism levied against the intelligence community seemed wholly deserved, especially after a mosaic of terrorist warnings contained in neglected FBI field reports came to light. But the criticism should have been tempered. It was neither realistic nor fair. The seeming understatement of the risk of foreign terrorism inside U.S. territory once again can be characterized as a reasoned view. A logical analyst would not have transcended the rules of evidence and could not have divined the intentions of the terrorists.

To illustrate this case, assume that the top analyst (or leader) initially estimated the risk of an attack on the United States by a terrorist group flying hijacked planes to be one-tenth of 1 percent. Then how much should the expectation of attack have grown after receiving, say, four successive intelligence reports warning of an imminent attack? The surprising answer based upon the rules of logic, and assuming the same error rates used in earlier calculations (25 percent rate of failing to detect an attack that is actually underway; and 25 percent false alarm rate) is that the probability would grow from less than 1 percent to less than 10 percent after four alarming reports in a row (see Figure 3).

Once again, this does not suggest dogmatism in the face of discrepant information. On the contrary, it shows that a belief should not be overridden lightly. The math shows that a person whose initial expectation of a terrorist attack is very low will need to be exposed to a stream of alarming evidence – seven intelligence alarms in a row – before the person logically should estimate the risk of attack to exceed 50 percent.

[Figure 3]

Initial and Revised Expectations of Terrorist Attack (Given Attack Warning) Assuming a Warning System with 25 Percent Types I and II Error Rates.

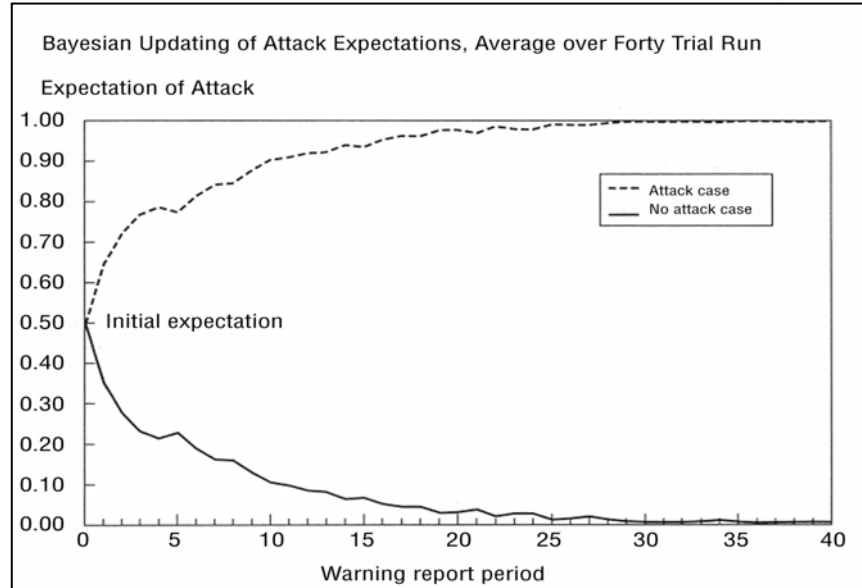
Initial estimate ^a	Revised estimate given attack warning															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0.0001	0.000+	0.001	0.003	0.008	0.024	0.068	0.179	0.396	0.663	0.855	0.947	0.982	0.994	0.998	0.999	1.000
0.001	0.003	0.009	0.026	0.075	0.196	0.422	0.686	0.868	0.952	0.983	0.994	0.998	0.999			
0.01	0.029	0.083	0.214	0.450	0.711	0.880	0.957	0.985	0.995	0.998	0.999					
0.05	0.136	0.321	0.587	0.810	0.927	0.975	0.991	0.997	0.999	1.000						
0.10	0.250	0.500	0.750	0.900	0.964	0.988	0.996	0.999								
0.20	0.429	0.692	0.871	0.953	0.984	0.995	0.998	0.999								
0.30	0.563	0.794	0.920	0.972	0.990	0.997	0.999									
0.40	0.667	0.857	0.947	0.982	0.994	0.998	0.999									
0.50	0.750	0.900	0.964	0.988	0.996	0.998										
0.60	0.818	0.931	0.976	0.992	0.997	0.999										
0.70	0.875	0.955	0.984	0.995	0.998	0.999										
0.80	0.923	0.973	0.991	0.997	0.999											
0.90	0.964	0.988	0.996	0.999												
0.95	0.983	0.994	0.998	0.999												
0.99	0.997	0.999														
0.999	1.000															

a. Degree of belief in the hypothesis "a terrorist attack is under way."

This slow revision of subjective opinion eventually converges on objective reality (see **Figure 4**) which illustrates a case in which the initial estimate is 50 percent). As more intelligence data become available and are brought to bear on opinion, the weight of initial opinion declines, eventually yielding completely to the data - assuming the data are not intentionally twisted or manufactured for political reasons.

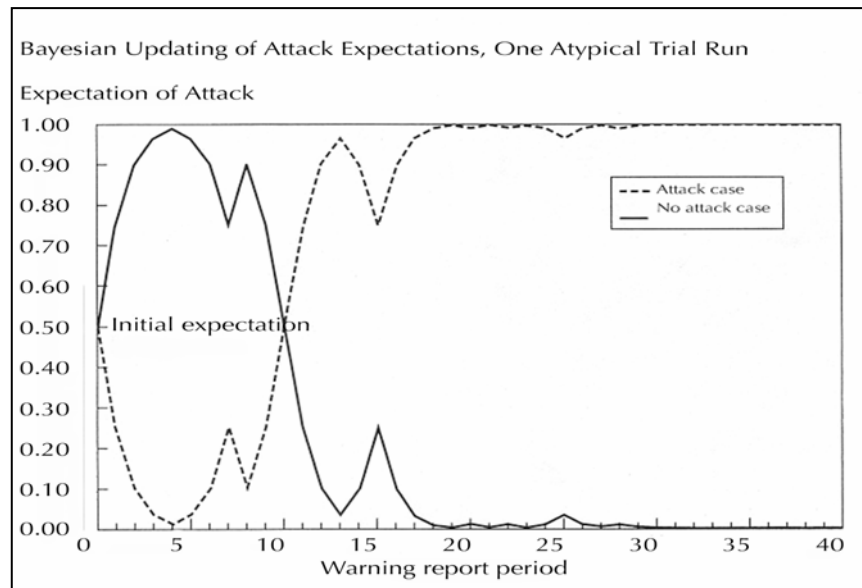
How long does Bayes' formula suggest it should take for this process to iterate itself to the truth? Unless some momentous event like an actual terrorist strike or the actual use of mass-destruction weapons intrude to compress the iteration time, 10 to 20 successive cycles of judgment are normally necessary across a fairly wide spectrum of conditions. Over the course of these cycles of assessment and warning there would be, in the case of an actual attack underway, occasional failures to detect the attack (reflecting a 25 percent error rate) which in turn stretches out the period of warning review needed to reach the proper conclusion. By the same token, in the case of no attack underway, occasional false warnings (reflecting a 25 percent false alarm error rate) would stretch out the time needed to realize that no attack was actually being mounted. A computer simulation was run to capture these statistical risks in which erroneous warnings would be mixed in with correct warnings (which the intelligence collection achieves 75 percent of the time).

[Figure 4]



In short, anything less than a lengthy series of spy reviews would represent a rush to judgment. Bayesian calculations in fact show that it is quite possible for the intelligence findings to be wildly off the mark for 10 or more cycles of assessment before settling down and converging on the truth (see Figure 5). A run of bad luck – failures to detect an actual attack, or false alarms if there is no actual attack – could drive the interpretation perilously close to a high-confidence wrong judgment. Although it would be unusual to experience a long run of bad luck, it is probable enough to play it safe and not preemptively attack or adopt draconian homeland defense measures after only a few intelligence reports in succession have set alarm bells ringing loudly.

[Figure 5]



CONCLUSION

This perspective on the intelligence process leads to an exonerating statement and a cautionary note. The exonerating point is that people who clung to their belief that Iraq possessed mass-destruction weapons in spite of the inability of intelligence efforts and inspectors to find them during the run up to the 2003 invasion, and even people who still believe today that mass-destruction weapons remain hidden in Iraq, have had a strong ally in logical reasoning for a lengthy period of time. A case can be made that their view has been intellectually the most coherent and consistent view of the threat. However, logical minds open to fresh intelligence reports should by now harbor serious doubt. The facts on the ground are speaking loudly for themselves in challenging the presumption used to justify the war with Iraq.

The cautionary note is that Bayesian math points to a fairly slow learning curve that also challenges the wisdom of making preemption a cornerstone of U.S. security strategy. The intelligence burden of this strategy is generally very heavy, too heavy for any leader to consistently shoulder. In all likelihood, a prudent interpretation of intelligence would fail to clarify the actual threat, the appropriate targets, and other contours of a preemptive strike. The strategy is not a feasible or sensible approach to U.S. national security.

Bayesian analysis proves that even good intelligence and interpretation are unlikely to meet the high threshold of waging preemptive or preventative war. In reality, intelligence information is more murky than our Bayesian analysis assumed. Bits of information in the real world are often ambiguous in their very meaning – thus two observers with different preexisting beliefs will often believe that the same bit of behavior confirms their beliefs – hawks seeing aggressive behavior and doves seeing evidence of conciliatory behavior.

Bayesian analysis does not confuse the meaning of bits of information, as though drawing balls of different color from a jar. And still, it shows what a mountain of evidence is needed to rationally change one's mind and arrive at the truth.

Bruce Blair is President of the Center for Defense Information in Washington D.C., and a former Senior Fellow at the Brookings Institution, Project Director at the Office of Technology Assessment, and Minuteman launch control officer in the Strategic Air Command. He holds a Ph.D. in Operations Research from Yale, has taught security studies as a visiting professor at Yale and Princeton, and was awarded a MacArthur Fellowship Prize in 1999 for his work on "de-alerting" nuclear forces.

*This paper was presented by the author at the 10th International Castiglioncello Conference “Unilateral Actions and Military Interventions: The Future of Non-Proliferation,” Sept. 18-21, 2003, Castiglioncello, Italy. The author is grateful to the Italian Union of Scientists for Disarmament and Professor Nicola Cufaro Petroni for comments on the paper. It draws heavily on Bruce G. Blair, *The Logic of Accidental Nuclear War* (Brookings, 1993). The author wishes to thank Rob Litwak and Robert Jervis for helpful and insightful comments. The author is solely responsible for any errors.

www.cdi.org