

Assessing the Value of Intelligence: Lessons for Leaders

Table of Contents

Foreword	4
Analyzing Investments in Intelligence Capabilities	6
The Changing Nature of Conflict and Intelligence	6
Responding to the Changing Nature of Conflict	7
A New Approach to Assessing Intelligence Value	7
Summary of Lessons Learned in Applying OR-Based Methods to Intelligence Investment Decisions.....	9
Lessons Learned in Leading Operations Assessments of Intelligence	10
Case Studies Assessing Intelligence Performance and Value	23
Introduction	23
Assessing the Need for Predator Unmanned Aerial Systems: 2006	24
Assessing the Intelligence Needs of a Special Operations Task Force: 2007	27
Assessing Conventional Ground Forces' Use of FMV: Late 2007	32
Assessing Armed ISR in a Counter-Indirect Fires Mission: 2008	35
Analyzing the Value of GMTI: 2008 through 2010.....	39
About the Authors	44

Foreword

On behalf of the IBM Center for the Business of Government, we are pleased to present this special report, *Assessing the Value of Intelligence: Lessons for Leaders*, by Chris Whitlock and Frank Strickland.

Applying power in all its forms to secure the present and future is ultimately a leadership challenge. That challenge is especially complex in the current century when the forms and patterns of security are changing in so many ways at an accelerated pace than ever before. The capabilities required to threaten a nation, region, or even global stability are available to both rich and impoverished nation states, as well as small networks of people who can and do operate relatively independent of any nation state. There is more data available than ever before to make sense of this era. However, leaders are in great need of capabilities that turn this data into knowledge that informs their discernment of the security strategies required in these dynamic and uncertain times.

To assist leaders with this challenge, the Center has embarked on a series of research activities focused on security, power, and leadership in a new century. This series includes radio interviews, articles, and special reports (For the first article in this series and links to the radio interviews, see: <http://www.businessofgovernment.org/article/power-security-and-leadership-21st-century>.)



Frank B. Strickland Jr.



Kevin Green

The following special report contributes to this series in two unique ways. First, it addresses lessons that leaders can apply to the management of intelligence capabilities. Much is often said, and rightly so, about the need for capabilities to turn data into knowledge or intelligence. The capabilities required to manage the intelligence enterprise receives far less attention. This special report provides five practical lessons that senior leaders can apply to assess the value of intelligence, thus managing both the operational and fiscal resources required to create intelligence.

Second, a large body of classified assessments of the value of intelligence is the basis for these lessons. As the reader will notice in the report, a number of national security leaders have commended these assessments. Thus, the lessons are grounded in a rich empirical, rather than conceptual, basis.

Leaders face difficult choices every day in allocating scarce intelligence resources to a complex array of global threats. Additionally, leaders must make wise investments of constrained fiscal resources to produce the intelligence capabilities required in an uncertain future. We believe leaders must demand that data and proven analytical methods inform these choices. The lessons in this report—perhaps the first of its kind—should help leaders establish a culture wherein data-driven assessments of intelligence value are the institutional norm. In doing so, they will improve the value of intelligence to power and security.

Frank B. Strickland, Jr.
Senior Fellow
IBM Center for The Business of Government

Kevin Green
Vice President
IBM Federal

Analyzing Investments in Intelligence Capabilities

The changing military force postures in Iraq, Afghanistan, and other parts of the world, combined with historic budget deficits, have caused the U.S. Department of Defense (DoD) to assess the capabilities required for current and future military operations. This requires DoD to make difficult investment decisions among many priorities. Investment decisions in the intelligence portfolio are particularly important as intelligence is the basis of all security plans and operations, including those intended to deter war.

In the past six years, a body of operations research (OR) assessments has proven helpful in informing intelligence investment decisions with hard, quanti-

tative data on the value of intelligence. Leaders can apply the lessons from these OR assessments to the challenges currently facing them in making investment decisions across various needs and capabilities in the intelligence portfolio.

*“If we have the intelligence advantage,
we can win.”*

— LTG Stanley McChrystal
May, 2007

The Changing Nature of Conflict and Intelligence

In 2006 the Iraq Study Group defined the situation in Iraq as “grave and deteriorating.” By year’s end the president had begun major shifts in leadership and strategy. Since the end of the Cold War, the U.S. defense community had primarily thought about and planned for major mechanized conflicts similar to the Desert Storm operation in 1991. While the U.S. military had engaged in some irregular warfare operations, e.g., in Somalia and Bosnia, neither U.S. warfare doctrine nor material acquisition focused on or contemplated major counterinsurgency and counterterrorism campaigns. Intelligence systems were equally ill-prepared for the changing nature of conflict, having

progressed from Desert Storm with relatively minor modifications to systems largely designed for a Cold War-mission environment. The scope and scale of the Iraq insurgency were greater than anything experienced since the Vietnam war. Intelligence capabilities were not initially prepared to distinguish insurgents from civilians and to provide the information needed to answer questions about individuals, tribes, and human networks.


Responding to the Changing Nature of Conflict

Faced with a fundamentally different threat environment, U.S. forces in Iraq and Afghanistan made urgent requests for large quantities and various types of intelligence systems. For example, the requests for increased full motion video (FMV) by U.S. forces seemed insatiable. As requests soared for intelligence capabilities, DoD worked to identify solutions and quickly deploy those capabilities into the two theaters of operation. Neither DoD nor the intelligence community at large had anticipated the pace or scale of such requests. As a result, a flurry of efforts was made to expand the capacity of existing capabilities, field new quick-reaction capabilities, and explore innovative applications of existing and new technologies. Not all of these would work equally well, nor were all equally scalable. The situation in Iraq drove an imperative to quickly field as many solutions as possible. At the same time, the secretary of defense and DoD leaders craved an understanding of the solutions' operational value, so that they could focus resources and energy on the solutions with the most value.

A New Approach to Assessing Intelligence Value

In 2006, the Office of the Secretary of Defense (OSD) and Joint Staff cosponsored what became a series of operations research (OR)-based

assessments to quantify the performance of intelligence capabilities and inform critical investment decisions for improving the mission value of intelligence in the war zones. These assessments began with a focus on the need for U.S. Air Force Predator unmanned aerial



"If you have not seen the OSD HVI analysis, you need to. This is how we should be doing our work to identify and prioritize intelligence needs. We should not be doing these 'split a dollar drills.'"

— MG Mike Flynn, CENTCOM, J-2
July 2007



systems, but rapidly expanded to encompass virtually all of the intelligence capabilities involved in the war effort against multiple military missions. The assessment findings gave DoD principals (such as the under secretary of defense for intelligence and the vice chairman of the Joint Chiefs of Staff) an ability to responsively cut through the complexity of intelligence investment deci-

sions with insights derived from hard performance data.

While the discipline of intelligence analysis relies on data about threats and the environment, assessments of the value of intelligence capabilities have traditionally relied on anecdotes, surveys of subject matter experts, and other qualitative data, along with some modeling and simulation. These new OR-based assessments were unique in that large volumes of intelligence outputs were tested against data from actual military operations to determine mission value. Teams gathered the intelligence and operations data using analytic techniques and tools, such as classified network crawlers, text parsers, and processes to format and store the data for analysis. The teams then applied advanced analytics to test whether the intelligence outputs correlated to operational missions. In addition to determining value, these tests also illuminated the root causes of performance shortfalls and opportunities for improvements. While the approach was focused on empirical data from intelligence and operational systems, consultants traveled into theater to observe the end-to-end intelligence cycle—tasking, collection, analysis, and communication—and how intelligence integrated with operations. Interviews, direct observation, and surveys were important supplementary data, but the approach was anchored on quantitative data from the actual operations. The resulting assessment approach was a blend of OR and commercial consulting methods, applied through an understanding of military operational domain.

In 2008, the Intelligence, Surveillance, and Reconnaissance (ISR) Task Force would take up the primary sponsorship of these assessments

with important co-sponsors including the Office of Cost Assessment and Program Evaluation (CAPE); Under Secretary of Defense for Acquisition, Technology, and Logistics; and the Joint Staff J2 and J8. General James Cartwright, then-vice chairman of the Joint Chiefs of Staff, became the most avid user of the findings derived from these assessments. In commenting on the use of these assessments for intelligence budget decisions, General Cartwright would say, "...the work that we've done with operational research analysts out in the field on our ISR systems ... make this as quantitative as ever I have seen in one of these budget developments."

OR-based assessments of intelligence value can help leaders address the current investment decisions facing those who manage the intelligence portfolio. This report provides a brief discussion of five lessons learned that leaders can use in applying OR-based methods. The report also provides an unclassified summary of three of the assessments.

Summary of Lessons Learned in Applying OR-Based Methods to Intelligence Investment Decisions

1. Analysis must get inside formal requirements to understand the mission value of intelligence, trade-offs among priorities, and alternative solutions.
2. Quantifying intelligence data and information is a necessary step to understanding the relationship between intelligence outputs and mission outcomes.
3. The performance shortfalls of current capabilities should be quantified, but assessments must focus on solutions from the very beginning of the project.
4. The people and organizations involved should be prepared for constructive conflict.
5. OR-based assessments are difficult to execute, but the results provide better information to base difficult decisions on than alternative methods do.

Lessons Learned in Leading Operations Assessments of Intelligence

Over the past seven years, teams of consultants have performed 100 assessments of intelligence capabilities using an OR-based consulting approach the authors created, “Edge Methods.” These assessments have addressed a variety of missions including counterterrorism, counterinsurgency, counter-proliferation, counterintelligence, strike, and cyber. The unclassified case studies, summarized later in this paper, are a subset of these assessments focused on intelligence activities in the Iraq and Afghanistan war zones between 2006 and 2009.

Performing these assessments and communicating the results to some of the nation’s most senior defense and intelligence leaders provided a substantive and practical basis for deriving lessons learned. The five lessons presented here can assist leaders and others who want to drive improvements in the assessment of intelligence capabilities’ performance and value to the users of intelligence. The fiscal challenges and operational transformations of today make this a timely and important topic.

1 Analysis must get inside formal requirements to understand the mission value of intelligence, trade-offs among priorities, and alternative solutions

ISSUE

Intelligence officers and the ultimate users of intelligence—policymakers, commanders, operators, and system developers—employ various processes to submit requirements for intelligence. The difference in these requirements processes is a natural function of the intelligence enterprise’s complexity—roughly \$80 billion annual budget executed by 16 different departments and agencies, seeking to understand a set of security issues spanning nearly 300 countries and organizations, most of which go to great lengths to hide their true intentions toward America and her allies. Some departments and agencies, such as the Department of Defense, have well-established requirements processes;

others less so. Regardless, the intelligence enterprise receives all submitted requirements and tries to respond with operational capabilities, if possible, while also considering new capabilities as necessary to meet the requirements. This creates one obvious challenge, as the desire for intelligence always exceeds the capacity of the intelligence enterprise.



Understanding these many requirements is also complicated by several biases that affect how users state their requirements. Requirements for current intelligence tend to reflect a view of what is feasible or reasonable based on the intelligence officer's or user's understanding of current intelligence capabilities. When stating requirements for new intelligence capabilities, intelligence officers and users again tend to reflect their understanding of available technologies. As one of many examples, prior to the widespread use of FMV from unmanned airborne systems the documented requirements for motion imagery called for collection of large numbers of individual motion imagery, but each with very short duration. These requirements reflected the understanding of technology at that time and how that technology would be used. Once FMV capabilities proliferated during the war, the actual operational uses were often the complete opposite of the previously stated requirements. Motion imagery collection often occurred in relatively fewer numbers, but over tremendously long duration. The cognitive biases that affect requirements for intelligence are not unique to intelligence. They are well-established in research of customer behavior.

Perhaps the greatest issue with intelligence requirements is the user's tendency to state desired solutions instead of operational needs. This is especially true in requirements for intelligence collection. In other words, the user will demand a certain amount or number of specific collection systems instead of stating an operational problem and the associated intelligence needs. This is especially true in situations where deployed military forces are concerned, as the military commander will naturally seek to amass as much capability as possible given the grave consequences of their mission. Here again, the tendency for users to assert required solutions vice their needs is an established customer

behavior pattern, and civilian and military users of intelligence will regularly exhibit this behavior in stating their requirements.

IMPLICATIONS

Few requirements provide an understanding of the user's operational need; i.e., what the user is trying to accomplish and how intelligence specifically enables the operational objective. This limits the ability of the intelligence enterprise to consider alternative solutions to the need. Trying to understand a security threat in a dynamic, and often hostile, environment is one of the most difficult intellectual endeavors. It is therefore important to consider a range of solutions to the intelligence need. The breadth of technologies and tools in the intelligence enterprise is such that no one user, even one with substantial staff support, is likely to consider the full range of solutions. The lack of information on the actual need hampers the ability to consider a range of solutions and select the solution(s) that best meet the need.

A lack of understanding about the operational need and desired outcome also inhibits prioritization. The many different users of intelligence, the many topics that concern them, and the multitude of questions across those topics present difficult choices that require prioritization of intelligence resources. Even within a given country or region, the intelligence enterprise must make trade-offs among requirements. These trade-offs are best decided in close collaboration with the intelligence user. Requirements that lack information on the operational need inhibit both that dialogue and the responsive prioritization of resources.

Requirements statements, especially those exercised by the Department of Defense, also take on a high degree of bureaucratic formality. This further inhibits the analysis necessary to understand the operational need, consider alternative solutions, and prioritize among requirements. Headquarters personnel become hesitant to dig into the details behind a requirement as this is perceived as challenging the user's authority or the validity of the requirement. A requirement that has been



validated by a senior commander or high-level bureaucratic body takes on an almost canonical quality. This becomes a bureaucratic Catch-22 as most are hesitant to question the requirement even though it does not contain sufficient information on the need.

LESSONS

An organization's formal statements of intelligence requirements often mask the actual mission needs and value required by the organization's decisions and operations. This inhibits analysis of alternative capabilities that provide the greatest mission value, as well as analysis of priorities across many, often competing requirements. Headquarters staff and solutions providers must pay careful attention to formal requirements statements. At the same time they must drive analysis to understand the underlying mission needs and the value of the requested intelligence to the user's desired outcome or results. Analyzing mission needs and value behind the requirements, as well as analyzing alternative solutions and priorities among the needs, are essential responses to honoring the requirements' intent of improving the value of intelligence capabilities to a large and diverse set of users.

Leaders must establish a culture wherein analysis of requirements is part of honoring the user. Users are best served by an intelligence system that meets their needs, not one that reflexively responds to their stated requirements. In fact, the best response to the user's requirements is one in which the user's need for intelligence is satisfied without the user having to ask. Anticipating the user's need requires detailed analysis of the user's operations and how intelligence can best enable the desired outcomes. Such analysis also provides the basis for a productive collaboration between intelligence provider and user.

When analysis over time reveals a new or greater understanding of the needs for intelligence, it is important to use such analysis in evaluating intelligence programs of record. There is a bureaucratic tyranny that sustains programs, especially large programs, once they are formally established in the budget process. The targets of intelligence—the threats that intelligence is supposed to understand—are constantly changing, as is the environment in which they operate. Intelligence systems must be agile to keep up with these dynamics. This requires leaders to ensure that programs of record do not use historically validated requirements to fend off necessary changes demanded by current analysis of operational needs and benefits.

2 Quantifying intelligence data and information is a necessary step to understanding the relationship between intelligence outputs and mission outcomes

ISSUE

Some senior intelligence officers have argued that the performance or value of intelligence capabilities cannot be properly measured. This line of thought posits that intelligence is entirely art and defies any systemic assessment. This conflicts with the reality that the intelligence officer's mission relies heavily on data analysis to reach judgments about threats to security. However, the majority of intelligence officers are not scientists or engineers. Thus, the use of data, especially a variety of quantitative data, is not an entirely comfortable domain.

It is also fair to note that performance data itself will not create conclusions on the value of intelligence. Those judgments must come from the minds of consultants and ultimately from the decision-makers. That said, intelligence officers should be comfortable with the relationship between performance data and conclusions on the value of intelligence, as it closely parallels their work of intelligence analysis.

The misuse of quantitative data in assessing intelligence value has given some intelligence officers justifiable concerns. Some assessments begin and end with simple tallies of collection reporting or some other measure that either does not immediately reveal value or is altogether irrelevant to an assessment's questions. These volumetric statistics, as some call them, are starting to proliferate given desktop computing capabilities, greater sharing of intelligence data, and increasing calls from oversight bodies for performance measures.

Some objections to quantitative assessments may be based on programmatic and political factors as well. Any intelligence capability that has a sizable budget will naturally accrue a set of advocates in government oversight organizations—both executive and legislative branches—in the user base and in industry. In some cases experienced senior officials object to the mere fact of an assessment, regardless of the methodology. Deployed military commanders can, for example, object to someone in Washington seeming to question their point of view on an intelligence capability's value. These objections further inhibit the use of performance data to bring the contribution of intelligence capabilities into the light of resource decision-making processes.

IMPLICATION

What is the alternative to using performance data to assess intelligence value? If it is an expert's judgment alone, which expert can be relied on in an enterprise as large and with as many crosscutting responsibilities as the intelligence community?

How do we understand what drives performance and value unless we rigorously assess them with data? Unfortunately, the value judgment often defaults to the senior official in the room, making the decision with the information available. Historically, the performance information available has come from the decision-maker's experience and performance anecdotes put together by various



constituencies for and against a capability. The combined implications of these issues can lead to value judgments being made purely in the context of the budget process, which almost ensures perpetuation of the status quo. Painful choices and changes to the baseline operations or programs are inherently difficult, and almost impossible without a strong data-driven case for action.

Not every quantitative assessment will yield significant conclusions. This reality, and the intelligence officer's relative lack of familiarity with quantitative data, should not inhibit the use of quantitative data. The misuse of measurement should not dissuade leaders from using quantitative data to assess intelligence value any more than the occasional mistaken use of intelligence in national security should dissuade us from collecting and analyzing intelligence.

LESSONS

Useful insights on intelligence value are not obtained by simply measuring the volume of intelligence collected and counting the number of intelligence reports relative to the requirements. However, quantifying relationships between inputs (intelligence requests) and associated outputs (collected data and analyzed information) is a first step in the assessment process. From this baseline of fact-based operational data,

the user can explore such value-laden questions as: Given the priority needs in Country X, what is this capability contributing relative to other available capabilities? What causal factors are driving shortfalls in intelligence? Given the strengths and weaknesses of various capabilities, how could performance improve by changes to the intelligence posture? What are the trade-offs in executing these alternatives? Consistent with the stated intent of current Director of National Intelligence Jim Clapper, quantitative data enables characterization and analysis of the relationship between the requests for intelligence, intelligence outputs, and the user's desired outcomes.

As the case studies in this paper illustrate, quantitative analytics is a feasible and useful way to inform decisions on the value or contribution of intelligence to a user's outcome. The U.S. intelligence and national security enterprise is supported by many automated systems—systems rich in data on operations, intelligence collection, and analytical products. This data is accessible and useful to consultants with the right skill sets and tools—primarily commercial tools that do not require lengthy software development efforts before any action is taken. This data supports assessments of intelligence value to a wide range of military operations, as well as national issues. Large volumes of various quantitative data also support conclusions on the normal performance or value a user can expect from an intelligence capability.

3 The performance shortfalls of current capabilities should be quantified, but assessments must focus on solutions from the very beginning of the project

ISSUE

Operations assessments require a team to execute a wide range of difficult tasks, such as locating and gathering diverse data sets, extracting the data relevant to analysis, correcting formatting and other issues in the data, and performing other tasks leading to an analysis of intelligence data relevant to operational outcomes. In the midst of these complex analytics, it is easy for the team to become fixated on finding performance shortfalls. During the early operations assessments of intelligence in 2006–2007, the assessment teams had a tendency to focus narrowly on quantifying the shortfalls in intelligence value. The assessments quantified, for example, the actual performance of ground moving target indicator (GMTI) systems against high value individual (HVI) operations. The quantified performance was strikingly less than what one

would conclude from operational anecdotes. While the vice chairman of the Joint Chiefs of Staff, under secretary of defense for intelligence, and some operational commanders were glad to understand the actual performance of GMTI, they quickly began to press the assessment teams for solutions to improve performance.



IMPLICATION

Isolating performance shortfalls and their root causes are necessary steps in the assessment process, but leaders and users are interested in solutions, not shortfalls. Operations assessments of intelligence value must produce more than statistical plots and utility curves. If all a team does is characterize performance problems, then the team essentially puts the burden of improving performance on the decision-makers and users. While that may seem politically safe at times—the team avoids telling the customer what to do—it is far less helpful than developing an actionable point of view with steps to improve performance.

LESSONS

Leaders must ensure that the assessment team is focused on finding solutions to problems, not simply defining problems in exquisite quantitative detail. This begins at the very beginning of an assessment by focusing the team on solutions hypotheses as well as problem hypotheses. Formulating solutions hypotheses requires consultants on the team to possess expert knowledge of current and prospective capabilities and technologies relevant to the performance shortfall. Additionally, the team must understand the non-material improvement levers, such as training and process improvement, preventing a narrow and limited focus on just system solutions. At times the non-material solutions will enable performance improvements quickly and at relatively lower costs. As the assessment progresses and the team gains an increasingly data-driven understanding of the performance issues, the team constantly refines the solutions hypotheses into a set of actionable recommendations to improve performance.

4 The people and organizations involved should be prepared for constructive conflict

ISSUE

Quantifying the performance of intelligence capabilities often tests conventional wisdom and produces surprises, given that the value of intelligence has traditionally been determined by expert intuition and anecdotes. In an enterprise as complex as the intelligence community, the ability of any given expert to understand performance details across the many capabilities in the enterprise without robust performance analytics is close to impossible. Intelligence officers, commanders, and operators, engaged daily in creating and using intelligence will develop perceptions on the value of intelligence capabilities based on the situations they experience and their intuitive judgment on the value of intelligence. This is not only normative, but these judgments are important to consider in the assessment process. However, the fact is that intelligence officers and customers of intelligence are focused on their own operations, and rarely have the time, tools, and expertise to sort out



detailed performance parameters, causal factors, and solutions in the intelligence enterprise. Further complicating this are the political relationships between Washington headquarters and deployed organizations, as well as the inertia of funded programs of record. Government programs of record accumulate fierce constituencies in the states and even with deployed command-

ers and operators. Quantitative analysis and logically derived conclusions will not eliminate disagreements and conflicts, as no major decision is entirely rational. Once hard performance data and assessment findings are on the table, the probability for conflict is high.

IMPLICATION

Rather than invoke disagreement and bureaucratic conflict, formal decision-making processes in large bureaucracies will often resist, if not reject, content that seems to run counter to their senior leaders' perceived interests or preconceptions. Staff officers are conditioned to

view a good meeting as one in which consensus is reached without any disagreements. The prevailing culture is somewhat in response to the African proverb, “When elephants fight, it is the grass that suffers.” There is a natural and understandable inclination to accept deployed commanders’ positions at face value. Passive agreement with the customer, even when that customer is a senior commander, is not synonymous with serving the customer’s needs.

LESSONS

Senior leaders in the DoD and intelligence communities must clearly signal their personal commitment to constructive debate and differing points of view, fueled by OR-based assessments of value. In fact, assessments that fail to generate any firm reactions probably need close scrutiny of the assessments’ value. Senior leadership’s commitment gives its staff the bureaucratic courage required to move through periodic disagreements without backing away from the assessment objectives. The record of the OSD and Joint Staff in the attached case studies is impressive in this respect. Sponsors of these assessments anticipated conflict and were prepared to deflect heat from customers of intelligence and occasionally redirect the assessment team. For its part, the assessment team must be expert in executive communications, change management, and bureaucratic savvy in order to manage these inevitable tensions to productive ends. Thus, we believe that practitioners of this approach must have the competencies of an analytical consultant vice simply those of an OR analyst.

5 OR-based assessments are difficult to execute, but the results provide better information to base difficult decisions on than alternative methods do

ISSUE

OR-based assessments, such as that applied in the attached case studies, are difficult to execute but quite feasible. Success requires a unique blend of expertise in operations research, customers’ missions, intelligence solutions, and consultative skills.

In contrast, it is much easier and faster to interview or survey a range of experts, either individually or through some type of focus group, and synthesize the insights from this information. A class of software capabilities exists to assist in structuring the questions and capturing the experts’

ideas. Some of this software will even enable quantification of experts' judgments on the value of intelligence or an intelligence capability. In addition to the relative ease of implementation, relying on a group of experts may also provide some political cover for the results, especially if one or more of the experts are highly regarded by the decision-maker(s).

Alternatively, computer-based simulations of intelligence can generate detailed quantitative data on how intelligence systems perform against a given target scenario, but they also present a number of complications. Such simulations are based on models of the intelligence capabilities, targets, and environment. Computer programmers create these models, describing not only an object's attributes, but its notional behaviors or operations as well. For example, a satellite collection capability will be represented in a model that describes the satellite's orbit, the associated sensors on the satellite, and rules governing when and how the satellite's sensors can collect against a target on the earth. These models give analysts the flexibility to change one or more elements of a scenario while holding the other elements constant, and thus generate performance data on variety of conditions. However, if the many variables associated with most conflict scenarios—such as the number and variety of intelligence and operations systems, and the variations in the operating environment—are considered, the complexity of simply constructing a modeling and simulation-based approach to assessments is apparent.

IMPLICATION

Subject matter experts and computer-based simulations both have distinct advantages and disadvantages in assessing intelligence value. The expert-based approach is fast, and experts often have valuable observations on how operations actually work. Conversely, research has established the biases that distort human observations, even those by experts, especially in a topic as complex and multi-dimensional as enterprise-level intelligence performance. Computer simulations generate actual performance data from which conclusions can be drawn, and the computer models allow certain attributes of a scenario to be varied, while others are controlled. However, the performance parameters governing these scenarios are often taken from systems' technical specifications and interviews with experts. These may not reflect reality well and often substantially simplify how systems actually perform in the real world, introducing potential errors into the simulation's output. In addition to requiring substantial time to construct a simulation

involving multiple systems, the results often appear to come from a “black box” as the simulation itself is so complex that decision-makers cannot understand it.

LESSONS

OR-based assessments are uniquely placed among other approaches. Commercial software capabilities—driven by the explosion of business analytics in the private sector—enable consultants to responsively perform operations assessments, delivering initial results in as little as 60 to 90 days and sometimes faster depending on the problem. This approach produces a rich set of naturally quantitative data on the intelligence output and its effects on operations. Compared with subject-matter experts, this approach provides a stronger, substantive basis of quantitative facts from which to draw conclusions and to make and defend resource decisions. The quantitative data also supports statistical and other examinations of the causal factors affecting performance, as well as determining what performance levels are normative; i.e. those customers can rely on. At the same time, the operations assessment approach does include experts’ input, using it early in the project to frame performance hypotheses that are then tested with hard data.

OR-based assessments are generally much more responsive than modeling and simulation, as there is not a lengthy period required to develop and test computer models. Significantly, operations analysis is an ideal precursor to constructing models and simulations.

Programmers can use the performance data and statistics derived from operations analysis to construct models that are

faithful to real-world performance. An even greater lesson, perhaps, is that operations analysis helps narrow the range of problems for which modeling and simulation are required. This leads to the construction of more narrowly scoped models and simulations, which in turn take less time and fewer resources to construct and have less room for errors in the performance variables.



Finally, before launching a performance metrics effort or buying a performance dashboard, leaders should require an OR-based assessment that enables an understanding of performance. The analysis then illuminates what data and metrics are relevant, and if and how the organization can derive value from regularized reporting mechanisms such as dashboards. The commercial analytics software that enables responsive and affordable OR-based assessments also paves the way for data-driven assessments to become an integrated part of the enterprise (as detailed in the article, *Empirically-based Intelligence Management — Using Operations Research to Improve Programmatic Decisionmaking*). OR-based assessments are not a commodity service, but the returns in operational performance improvements and efficient use of resources justify the investment required in the assessments.

Case Studies Assessing Intelligence Performance and Value

Introduction

On April 6, 2009, Secretary of Defense Robert Gates gave an extended press briefing to explain the fiscal year 2010 budget and major changes in defense priorities. The timing of this briefing was unusual in that the secretary was announcing budget decisions and priority changes well before the formal defense budget was to be submitted to the president (usually late in the fall of each year). Chairman of the Joint Chiefs, Admiral Mike Mullen, was out of the country. Vice Chairman of the Joint Chiefs of Staff (VCJCS), General James “Hoss” Cartwright, joined the secretary for the briefing. The VCJCS chairs the department’s joint requirements process and works closely with the deputy secretary of defense to manage the department’s portfolio of capabilities.

A key part of the secretary’s presentation focused on rebalancing defense capabilities to institutionalize irregular warfare capabilities. This rebalancing of the portfolio included substantial changes to the intelligence capabilities required by irregular threats such as terrorists and insurgents. During the question-and-answer period, a reporter asked: “Can you tell us a little bit more, Mr. Secretary, about the analysis that went into these decisions? Even over the weekend there was some criticism that such bold decisions before the Quadrennial Defense Review, before this top-to-bottom review, perhaps don’t have the analytical framework that would be required. Can you give us sort of the 1-2 about how this all was put together?” In the response, General Cartwright said,

“On the intelligence side, the work that we’ve done with operational research analysts out in the field on our ISR systems, not just the platforms but how we move data and how we inform warfighters inside of decision cycles, ***these analytic pieces make this as quantitative as ever I have seen in one of these budget developments.***” (Emphasis added)

Among these “analytic pieces” referenced by General Cartwright was a body of assessments the authors’ teams performed beginning in 2006. General Cartwright was one of the executive sponsors of these assessments. In a series of meetings, often in his office on Saturday mornings, General Cartwright provided guidance on the key assessment questions and received detailed briefings on the assessment results. The case examples summarized here help illustrate the five lessons learned. The lessons learned were derived from a body of 100 classified assessments over the past seven years.

Assessing the Need for Predator Unmanned Aerial Systems: 2006

The Decision Problem

In 2006, the Air Force’s Predator unmanned aerial system (UAS) had become one of the primary FMV collection platforms. At that time, the Air Force had seven Predator systems in operation with an approved plan to acquire a total of 21 systems. Each Predator system reportedly provided FMV collection 24 hours a day, seven days a week, through the use of four unmanned aerial vehicles (UAVs). Four Predator UAVs enabled an operations concept that kept one UAV airborne collecting FMV at all times.

U.S. military forces operating in Iraq and Afghanistan were submitting requests for increased amounts of FMV, expressed as an aggregate number of hours of video per day. These requests went to the U.S. Central Command, which judged the requests and then submitted requirements to the Department of Defense through established processes. The number of FMV hours per day had been steadily increasing over time, causing senior executives in the Office of the Secretary of Defense (OSD) to question how much FMV was enough.

OSD officials were concerned about this issue in several respects. First, they wanted to anticipate the FMV demand so



that the department could plan accordingly and deliver the capabilities. In doing so, however, OSD leaders were troubled by the seeming lack of analysis behind the ever-increasing requests. The under secretary of defense for intelligence (USDI) in particular was troubled that analyses of the FMV needs were limited to largely qualitative approaches, such as those relying on subject matter experts' judgment. The department was trying to make difficult decisions on where to apply resources given a wide range of operational needs. In just the intelligence portfolio, for example, there were competing needs, such as increasing demands for ground moving target indicator (GMTI) radar. Some senior officials embraced a "wedding cake" collection strategy with GMTI providing the foundational capability and FMV acting as a complementary capability.

Summary of the Analysis

A deputy USDI, Mr. Tom Behling, commissioned a team of consultants to quantitatively assess the FMV need. The team initially built understanding of the need by reading secondary source documents and interviewing operational experts, many of who had flown Predator UAVs in the war zones. From this understanding the team formed a hypothesis that decentralized basing of the Predator systems might enable each system to collect greater numbers of FMV hours. The hypothesis was influenced by the relatively slow flight speed of the Predator and the assumed distribution of targets over a wide geographic area. Thus, multiple Predator bases, distributed near clusters of targets, would provide less transit time to the targets and more collection time.

In 2006 the Predator and UAS FMV were still relatively new capabilities. This presented an immediate problem in that the archiving of data from the Predator's operations was still relatively immature. Fortunately, the team met a contractor engineer who had access to a Predator FMV archive, and she agreed to parse out the telemetry from the video. This telemetry, or metadata, provided an essential source of empirical evidence on how the Predator UAVs, and their FMV sensors, were actually operated in the war zones. There was no metadata catalogue, at least that the team could find, so the consultants went to work figuring out the metadata's structure and relevance to the assessment.

The team also gained access to a repository of significant activity reports, "SIGACTs", on the activities of insurgents and U.S. forces. This also provided an essential empirical source of data on what was actually happening on the ground in terms of attacks on U.S. forces and

U.S. forces' operations. The team created computational techniques in statistical software as well as in the geospatial analysis system, ArcGIS, to analyze Predator UAVs' performance in context of insurgent attacks and U.S. operations on the ground. While accessing the FMV metadata and SIGACTs data was essential to the assessment, the team had to conceive and create techniques to clean millions of records before the analysis could even begin. After the data was rendered in a form useful to analysis, then the team created additional techniques to quantify the relationships between events on the ground and FMV collection.

Summary of the Results and Implications

Analysis of the data demonstrated that the team's hypothesis for distributed bases was invalid in the war zones. The targets were not distributed such that a distributed basing concept would produce more FMV hours per UAV.

The team's analysis also uncovered a major error in the Air Force's documented assertion that one Predator system provided 24x7 FMV collection. In fact, the data demonstrated that each Predator system was delivering much less than 24x7 collection. The difference was so substantial that the team initially thought there were errors in the data. However, the team learned that a limitation in the Predator ground system prevented the system from delivering a 24x7 operational orbit, even though a 24x7 orbit was the assumed planning factor throughout the OSD and Joint Staff. The team constructed a model that demonstrated how changes in the ground systems, and some changes in target prioritization, could generate substantial increases in the aggregate number of FMV hours. In fact, the model demonstrated the potential to meet the increased demand for FMV hours with a range from 13 to 17 systems, vice the 21 systems planned, potentially freeing some funds for investment in other capabilities.

While the ground station finding was useful, analysis of the war zone FMV data and other motion imagery requirements held by the National Geospatial-Intelligence Agency (NGA) caused the team to begin questioning whether aggregate FMV hours was a useful measure of the operational need. Aggregating the total number of FMV hours required in a day obscured important data about the operational need, such as the amount of time FMV would spend on a target and the distribution of targets over a geographic space.


The ink was hardly dry on this assessment when the Pentagon received a request from Special Operations Forces (SOF) for the equivalent of 30 Predator systems just for SOF war zone operations. This remarkable request furthered the team's concern over measuring the operational need by the number of Predator orbits or aggregate FMV hours in a day.

Assessing the Intelligence Needs of a Special Operations Task Force: 2007

The Decision Problem

Among the 100 quantitative assessments of intelligence value conducted from 2006 to the present, a detailed assessment of ISR support to the Joint Special Operations Task Force (JSOTF) is an exemplary case that demonstrates the benefits of evidence-based operations research. In 2007, the Defense Department had to respond to a remarkable JSOTF request; a Joint Urgent Operational Need (JUON) for 30 medium-altitude FMV orbits to support the mission against al

Qaeda in Iraq. The JUON was remarkable because it amounted to a fourfold increase in the existing Predator fleet at a time when the Office of the USDI was recommending limiting the Predator program to 21 systems or less. If the JUON was approved, the Department would be grant-



Beginning with the Predator analysis, it became apparent over time that stated requirements often mask the underlying operational need, and fixation on the requirements can cause teams to miss critical drivers of capacity, quality, or supporting capabilities needed for operational success.

ing JSOTF the entire Predator fleet plus nine orbits to support just one mission area, special operations. Recognizing the gravity of the decision, the USDI, along with two additional elements in the OSD, sponsored an assessment of ISR support to High Value Individual (HVI) campaigns.

Summary of the Analysis

The resulting analysis successfully constructed a coherent assessment of the relative contributions of 13 intelligence capabilities to operational success. Starting in late January 2007, the team analyzed and associated millions of records generated by various ISR assets with data on



2,500 special operations raids against al Qaeda in Iraq. The team often had to collect this data from relatively unstructured sources (like SharePoint, shared folders, and network diagrams) spread across several disparate databases and product repositories. In addition to the bulk data-gathering phase of the study, the team also conducted direct observation tests, interviews, and focus groups.

The fundamental data set supporting the team's analysis was a classified catalogue of daily raids conducted by Special Operations Forces (SOF). This critical repository provided results of the raid, temporal and locational data of the engagement, some indication of what intelligence cued or tipped it (e.g., human intelligence, SIGINT, FMV, etc.), and other associated data. Although the intelligence cue or tip indicators were sketchy and could not inform findings, this data was helpful in developing hypotheses and analytics to test those hypotheses.

In addition to operational data, the team processed tens of millions of intelligence-related messages, reports, and data. Of all the bulk data processed, the highest volume data set was FMV telemetry: systemic observations created roughly every five seconds describing the position of the FMV platform, sensor parameters, and aim points. In addition, the team processed over 50,000 unstructured Tactical Interrogation Reports (TIRs) by creating scripts to extract all geographic coordinates, names, and other related data. Other data sources included collection management records, raid storyboards, network diagrams, source annotations, SIGINT target lists, FMV vehicle tracking files, document and media exploitation (DOMEX) records, GMTI products, and imagery.

In addition to parsing bulk data from structured and unstructured repositories, direct observation was essential to understanding the operational process known as F3EA (Find-Fix-Finish-Exploit-Analyze). First, find the target—meaning identify the individuals to be pursued and understand the general location or operations area. Next, fix the target in time and space, holding it under observation, until a force can engage it. Third, finish the operation, meaning the tactics, techniques

and procedures necessary to successfully execute the raid. Fourth, exploit captured documents and media; and finally, analyze the data with other intelligence to fuel and repeat the cycle.

Accordingly, the team visited operational locations to observe the JSOTF and supporting elements in action. The team observed the command and tactical leadership function but concentrated on the central operations center, which handled all detainees and DOMEX. It also directly observed the daily collection management process within the JSOTF to understand the appreciable trade-offs between operational objectives and available resources.

In concert with direct observation, the team also conducted interviews to add context and balance to the quantitative analysis. Thanks to the JSOTF's open and proactive leadership, the team was granted the freedom to probe any relevant aspect of the JSOTF's operation. This included multiple sessions with the JSOTF's commander and members of his staff. The Cryptologic Support Group as well as the HUMINT Operations Cell offered insight into the accomplishments and challenges of their respective operations. Interviews extended to domestic support as well.

Although the interviews, direct observation, and focus groups were constructive, the central element of the assessment was the empirical association of operational data with collected bulk data. To accomplish this, the team recruited members with diverse skill sets, primarily professionals with extensive backgrounds in intelligence operations, quantitative analysis, and advanced computing. The team created specialized computer scripts to parse the data and operational products (e.g., PowerPoint-based storyboards) to enable the analytics. Although the team leveraged commercial software, the tools used to analyze the data were largely limited to ArcGIS, custom scripts, and conventional software such as Microsoft Excel and Access.

The analysis focused on establishing spatial, temporal, and relational connections between the data and operational objectives of the raid. For example, all locational data derived from the collected intelligence data sets was geospatially plotted against each objective's location. Then the team placed a 100-meter buffer around all locational data (objectives and intelligence) to see what sources intersected with what objectives. Since all sensors have some degree of target location error and objectives were not always individual houses (they could be larger

compounds, for instance) the team performed excursions to test the sensitivity of the intersections as the buffers increased from 100 to 500 meters.

Although this analysis provided a basic understanding into how F3EA evolved, it did not provide adequate insight into the crucial “find” phase. For that, the team performed a range of temporal and relational tests by extracting all the names and locations produced by every detainee processed by the JSOTF. The team’s subsequent analysis assessing the relationship between the locations and targets steered successive analytics. As an example, assume “Abu Muhammad” was the target captured in a particular raid. The team not only wanted to know what sources provided the location, but also wanted to know what sources identified Abu Muhammad. This was typically something derived in SIGINT narrative reporting, SIGINT network analysis, interrogation reporting, or DOMEX. The team also was keenly interested in the temporal facets of this problem i.e., the sources that tended to lead in the identification. All the potential target names and locations produced by each detainee ever processed by the JSOTF. Then an analysis was performed to see how those locations and targets drove the targeting process going forward.

Summary of the Results and Implications

By May 2007, the team offered its initial position, which largely remained unchanged for the remainder of the study. In this initial assessment, intelligence capabilities were arrayed from top to bottom in tabular format indicating their relative contributions in the “find-fix-finish” phases of the F3EA cycle.

The analysis revealed two striking surprises. First, the impact of FMV was superior to all other intelligence capabilities in all phases of the process, even in the “find” phase, much to everyone’s surprise. Second, and perhaps even more surprising, GMTI was only a modest contributor. This was particularly unexpected because the department’s ISR strategy placed strong emphasis on GMTI; there were vocal constituencies for GMTI’s value to HVI campaigns; and the team had assumed that GMTI was a natural contributor to tracking vehicles. Other traditional ISR capabilities, such as satellite reconnaissance systems, also proved to be of only modest value in irregular warfare.

The team's ranking of ISR capabilities prompted some controversy. Programs and their associated constituencies are characteristically disposed to only see their capabilities in the most flattering light, creating the need for constructive conflict. The team subsequently sustained a protracted period of conflict—some constructive, some less so—as decision-makers reviewed the ranking and underlying performance data. For all the lower contributing capabilities, the team performed a root cause analysis to understand how the capabilities might improve their performance. This analytic considered information from direct observation, interviews, quantitative performance analysis, and input from technical experts. The team evaluated four causes for lower contribution:

1. Capacity—was performance low because we did not have enough?
2. Use—was this an issue with tactics, techniques, and procedures?
3. Modification—did the capability lack some specific feature that would impact performance (e.g., downlink)?
4. Phenomenology—did the basic sensing parameters apply well for the collection capability?

This 2007 OSD HVI analysis ultimately developed into the foundation of the department's understanding of intelligence performance against irregular targets. Very clearly, the resolution provided by any capability emerged as a dominant theme. This was true not only of spatial resolution but also temporal and relational. Continuous or near-continuous surveillance of a fleeting target was crucial to not only finding but also fixing the target to enable operational action. Identity-level resolution emerged as the grail for most of the collectors—positioning the operators to know with confidence that they were actioning the right people.

The results were timely and coincided with programmatic budgetary decision cycles. Drafted in four months, the initial assessment arrived in time for issue development at the Pentagon (that May) and concluded before the president's budget was finalized in December. In less than one year, the team completed its exhaustive study, furnishing USDI and OSD with robust statistical insight into the performance and relative value of each capability. Isolating what was under-invested with rigorous empirical analysis, the team improved the defense department's understanding of what capabilities impacted performance against irregular targets.

Assessing Conventional Ground Forces' Use of FMV: Late 2007

The Decision Problem

Warriors know firsthand the uncertainty and friction produced by war. It is difficult for anyone who was not in Iraq to fully grasp the complexity and difficulty of the situation, especially in the aftermath of the 2006 civil war. While the extraordinary physical danger and exertion of war are only experienced by those in the battle space, the effects of uncertainty and an unpredictable enemy also create problems for the Department of Defense at large as it develops strategy and equips forces for operations. In 2007 the department began a major shift in leadership and strategy of the Iraq war. During this time of tremendous change

"You've got to tell them this is [not smart]. Make it very simple."

— Admiral "Fox" Fallon,
Commander, Central Command
October 2007

and complexity, the military services struggled with what at times seemed to be an insatiable demand for intelligence by U.S. forces in Iraq. The U.S. force in Iraq was about 10 months into the surge. Perhaps because it takes some time to develop and deploy new intelligence capabilities, it was not uncommon to hear people argue about whether the department was building too much intelligence capacity. In the summer of 2007, Commander of U.S. Central Command Admiral Fallon requested an analysis of the conventional ground force's use of FMV. Admiral Fallon wanted to ensure that the increasing requests for FMV were necessary and, like some others in the department, he wanted to avoid providing more capacity than was required.


Summary of the Analysis

As the previous case study illustrates, an assessment team had already analyzed the value of FMV for SOF operations. However, the scope of conventional ground force's FMV needs was much greater than SOF in three respects: larger areas of operation to consider, a wider variety of missions, and a much larger and more diverse data set, driven by the number and variety of UAS supporting conventional forces. The team leveraged lessons from the SOF HVI assessment but needed to create an analytic that addressed the three factors of the conventional-force FMV assessment. The decision problem addressed not only the value

of FMV to various conventional force missions, but also the quantities and type(s) of UAS FMV required by the conventional force's missions.

One team member—a consultant with remarkably strong computational capabilities—conceived a new analytic technique for this problem. He envisioned processing the tens of millions of telemetry observations into operational activities. When the FMV camera moved its aim point more than a prescribed distance in one minute, this indicated that the operators were onto a new task or target. With the telemetry data processed in this way, the team was able to group and characterize all the telemetry associated with a particular activity. In some activities, the camera was staring at smaller targets for extended periods of time, up to many hours. In other activities, the camera was quickly searching very broad areas. The team also created a geospatial visualization of such that the activities appeared as dots in a geospatial information system, ArcGIS. Much more than a simple display, ArcGIS provides powerful computational capabilities. The team leveraged these in order to assess the temporal and spatial relationships between FMV activities and the mission activities reported by U.S. ground forces.

The team analyzed a variety of operational, geographic, and time-based views to understand the conventional ground units' use of FMV. These techniques, applied to months of FMV telemetry, helped to define the patterns in using FMV that would not be apparent to someone examining collection management documents or intelligence reporting. The



Virtually all of the war zone performance assessments caused some degree of conflict. On complex issues, especially when the consequences are substantial, conflict is to be expected once hard performance data is brought to light.

analysis characterized all the activities by what military unit was being supported, the length of the activities, and the amount of staring or searching in each activity. To understand the FMV's operational impact, the team was able

to correlate the FMV to operational outcomes through the use of spatial, temporal, and target relational tests between the FMV and the operational event. These novel techniques revealed surprise differences in the way SOF and conventional forces were using FMV.

Summary of the Results and Implications

SOF missions tend to have far fewer FMV activities, with each activity

occurring over a long contiguous collection period. This is consistent with the HVI stalking operations analyzed in the previous case.

Conversely, conventional forces tend to employ FMV on many more activities, each with dramatically shorter collection periods than SOF activities, and spread over large geographic areas. This pattern of operation is consistent with the use of FMV for target search or area scanning activities. As a specific mission example, the conventional forces would fly a UAV up and down a transportation route, trying to find improvised explosive devices (IEDs). An insurgent could emplace an IED in a short period of time, making the probability of detecting this activity over large areas, with an FMV camera, extremely low. Likewise, once an IED is in the ground it is extremely difficult to detect, especially from a stand-off technical sensing system such as FMV. On the rare occasion that FMV caught an insurgent emplacing an IED, or found a likely FMV already in place, it made for a captivating success story, demonstrated in the video. However, the probability of success was extremely low. Of all the IEDs that were found and cleared, soldiers and other people on the ground found the overwhelming majority. FMV was essentially a non-factor in finding IEDs although substantial resources were deployed for this purpose.

This finding led to a broader implication for the use of FMV and ISR. If military forces choose to use airborne FMV to support defensive force protection missions, then the amount of FMV and ISR required is essentially limitless. Military forces, whether on an operating base or maneuvering in the battle space, take defensive measures to protect against enemy attack. The forces are constantly moving, and are also vulnerable to a degree when on a combat base or post; thus, the need for defensive measures is nearly continuous. The payoff of airborne FMV as a defensive measure is extremely small due to the low probability of detecting a low signature threat just prior to an attack.

By contrast, offensive operations are focused in space and time, and deliberate their approach to tearing down the enemy's network. FMV and other forms of ISR can produce much better results for the resources committed in supporting offensive operations. The team found a tremendously higher success rate in SOF offensive operations as compared with those of conventional forces during the assessment period. The analysis indicated a strong connection between the amount of ISR dedicated to offensive operations and the operations payoff. Significantly, it may be that well-supported offensive operations over

time can make the force much safer from attack, suggesting that the best application of ISR for force protection is to focus the ISR on offensive operations.

This assessment initially met with great resistance from senior commanders in Iraq. Several factors were at work here, not the least of which was relatively poor coordination of the assessment between

U.S. headquarters and commanders in Iraq. The substantial tension did serve to put the issues of ISR allocation and needs near the top of everyone's priority list. The coordination issues were resolved, and an agreement reached for a broader conventional force ISR assessment in close collaboration with the units in Iraq. MG Mike Flynn, the U.S. CENTCOM J-2, wrote to the team in January of 2008 encouraging the next step:



“You guys did just what we needed you to do in the fall [2007]. While they were aggravated, that session put this ISR issue in plain view. [That OSD] study is the only thing I have ever seen that quantifies the value of ISR. Keep pressing full speed ahead ... I talked with the [Multi-National Forces—Iraq] intelligence leaders. They are excited about the study and are anxious to have you come out.”

As conventional force units rotated out of and into Iraq in 2008, the assessment's findings became useful to how the force's intelligence officers would use UAV FMV capabilities. For OSD and the ISR Task Force, this assessment was important to budget formulation and near-term acquisition initiatives for FMV and other ISR capabilities.

Assessing Armed ISR in a Counter-Indirect Fires Mission: 2008

The Decision Problem

In 2008, a team assessed ISR value and needs for a range of conventional force missions in Iraq. General Petraeus was actively involved

and directive in this process, shaping a set of case studies around several conventional force mission areas. Assessing the value of armed ISR in countering indirect fires (IDF) was one part of this body of assessments.

In March 2008, Shia militia in Sadr City launched a mortar-and-rocket campaign against the Green Zone, the area in Baghdad where the U.S. Embassy and Multi-National Forces—Iraq (MNFI) headquarters were located. This attack was likely a response to the Iraqi government's offensive in the predominately Shia city of Basra. In response to the mortar and rocket attacks on the Green Zone, coalition forces mounted a determined counterattack including manned and unmanned ISR and attack capabilities. MNFI leaders informed the team that an assessment of this operation would showcase the value of armed ISR.

During the planning, execution, and assessment of operations, military commanders receive an almost constant stream of briefings. These briefings obviously help shape perceptions about the operational value of capabilities. An armed-ISR platform, such as the Predator UAV, can be a powerful attack capability in that it can collect intelligence and attack enemies from the same platform. Moreover, some effects of a Predator attack are seen in the real-time FMV it collects. When watching the video of one or more of these attacks, it is somewhat natural to form a strong perception of the value of armed ISR.

Arming an ISR platform, however, comes at a cost. The additional weight of the weapons takes the place of fuel and/or the weight of more capable ISR sensor payloads. The UAV cannot fly as long due to less fuel, and a less capable sensor payload may be installed to

accommodate the additional weight of the weapons. The armed-ISR platform is clearly valuable for its multi-mission capabilities. However, the multi-mission capability diminishes the ISR performance. This assessment sought to quantify the utility of armed ISR in this mission. In this case, U.S. forces had a large number of manned aircraft and artillery for attack. There was a




shortfall of ISR capabilities. Understanding the value of armed ISR was important to decisions about the mix of armed ISR in the overall ISR portfolio.

Summary of the Analysis

As was typical of all these operations assessments, the team collected data through automated searches of the classified networks and in visits to units involved in the operation. The necessary data represented a wide variety of sources and data types, such as detailed logs from the counter-battery radars, data from the aviation weapons teams (helicopter gunship teams), and intelligence data and weapons data from unmanned aerial systems. This initial phase of this assessment presented a challenge, also typical to these operations assessments, to understand the various fields in a data record, cleanse any anomalies from the volume of data files, format the data such that analysis is possible, and database the data.

Assessing the value of ISR in every case is much more than quantifying intelligence outputs, however. The team must understand the data in context of the operational mission in order to analyze the relationships between the data (an output) and the operational outcome. Understanding the operational mission means understanding four primary elements: the terrain or environment in which the operation occurs; the threat's capabilities, behaviors, and desired objectives; U.S. forces' capabilities, operations, and objectives; and the interrelationships between these three factors.



Intelligence and operations systems generate an abundance of data. Understanding this data, including basic statistical characterizations, is necessary but insufficient to operations assessments. To understand the value of intelligence capabilities, consultants must assess the quantitative data in context of the customers' desired outcomes.

This counter-IDF operation was focused on a 4x5-kilometer area north-east of Baghdad, Sadr City, with a dense urban population. Sadr City was a violent area and home to many of the Shia militia. The team analyzed all of the data to understand how each capability was contributing to three phases of the operation: finding the mortar or rocket threats, fixing the threats in time and space, and finishing (successfully attacking) the mortar or rocket teams. The team's hypothesis was that

armed-ISR platforms play a leading role in all three phases, including the finishing phase. Analysis of the data, however, clearly demonstrated that the primary operational value actually came from the counter-battery radars providing very accurate positional data on the mortar or rocket launches, coupled with Army aviation weapons teams (attack helicopters) executing the attack or finishing phase. UAVs, both the US Army Shadow and the U.S. Air Force Predator, played a frequent role in fixing the target; i.e. holding the target under surveillance until an action could be taken.

However, armed ISR—UAVs with weapons—played a minor role in the finishing phase. The operational environment was a crowded urban air space. At that time, Balad Air Base, just north of Baghdad, was the busiest military air base in the world. Armed-ISR platforms would generally operate at altitudes above several other aviation platforms such as helicopters, some fixed-wing manned aircraft, and some UAVs. In other words, the armed Predators were flying well above several other aviation platforms in an airspace that was constrained by both the small size of Sadr City as well as the number of aircraft massed over this area. In order for the Predator to take a shot in the finish phase of the operation, the airspace underneath it would need to be confirmed clear. Civilians have some sense of the complexity and importance of air traffic control around a big civilian airport. Military forces in combat have not only the standard air management challenge, but must also be concerned with ensuring that fires from aircraft are properly controlled. There were some impressive success-story videos of armed Predators successfully attacking mortar-and-rocket teams. In the context of the overall operation, however, these successes were the exception, not the norm.

Summary of the Results and Implications

This analysis demonstrated that armed ISR was not always the killer app (no pun intended) that some perceived it to be. At that time, given the overall shortfall in ISR capabilities, it was more beneficial to the overall war effort to emphasize UAVs that were fully capable ISR platforms vice armed-ISR platforms. This conclusion was immediately helpful to defense department planners, such as the ISR Task Force.

The team also ran excursions of this assessment, looking at the patterns of mortar and rocket attacks across the country. As it turns out, the Sadr City challenge was atypical in that most of the events were

generated from a small area over a relatively short period. U.S. forces were able to saturate this area with radars, UAVs, and weapons. This enabled U.S. forces to conduct a very effective counter-IDF operation. Across the rest of the country, however, this approach would not scale well. This conclusion helped to place the value of armed ISR in the counter-IDF in a theater context instead of the one operation in an isolated area.


Analyzing the Value of GMTI: 2008 through 2010

The Decision Problem

The 2007 assessment of ISR's value to the HVI campaign created a good bit of conflict over the difference between the perceived versus actual value of GMTI. These conflicts have political and emotional elements, but they also have rational elements beyond the success-story phenomenon outlined in previous cases. In the case of GMTI, the department was making substantial investments in a number of

manned and unmanned airborne GMTI capabilities.

During the first Gulf War in 1990–1991, GMTI from an Air Force Joint Surveillance and Targeting Radar System (JSTARS) aircraft had



“We have to break the model of one system for one target”

— General James “Hoss” Cartwright
Vice Chairman, Joint Chiefs of Staff

proven itself as a game-changing capability against Iraqi armored and other vehicle columns. For example, JSTARS tracking of Iraqi vehicles was critical to the coalition's defeat of a surprise Iraqi ground attack in the battle of Khafji, the first major ground battle of that war. By 2007, the department faced major investment decisions to upgrade and maintain JSTARS as well as other funding for other GMTI platforms and technologies.

GMTI radars are part of a family of wide area motion imaging (WAMI) sensors. Other WAMI sensors collect literal video images, but over much wider areas than the FMV sensors. This family of WAMI sensors could potentially address two primary problems. First, while the available FMV sensors were of tremendous value—as noted in the previous assessments—the sensor field of view was very narrow. Looking at the ground through the standard FMV sensor, the operator could see a 250-meter area or less; this would be roughly one city block or less in a dense urban area such as Manhattan. Since insurgent and terrorist

activity could occur over much wider areas, and the FMV operator was not always certain where to look with great precision, there was a need for surveillance of wider areas.

Second, narrow field-of-view FMV also results in an ISR force structure of one sensor platform to one target. Counterinsurgency and counterterrorism operations would encounter situations wherein it was necessary to follow multiple vehicles or people leaving from a common departure point. Doing this with standard FMV meant dedicating one FMV platform to every target the operators wanted to follow. The FMV field of view was too narrow to keep more than one target in view, especially when the targets left in separate directions. As a force planning factor, one FMV platform for one target would not scale. WAMI technologies provided the potential for addressing both of these issues with corresponding benefits to not only operations, but also the amount and type of ISR force structure the department needed to acquire.

From 2008 through 2010, consulting teams completed three primary assessments of GMTI. These assessments considered GMTI's value to conventional force operations in Iraq, counterinsurgency operations in Afghanistan, and a set of actions stated in a memorandum from the Joint Requirements Oversight Council (JROCM). Several supporting assessments were also conducted during this time, such as an assessment of the GMTI collection requirements of one Regional Command headquarters in Afghanistan. This case study summarizes some of the analytics and results from these cases as a whole.

Summary of the Analysis

An OR-based assessment of intelligence will generally gather data to characterize the collection tasking as one of the first analytics. In these GMTI assessments, teams extracted data on tens of thousands of collection requirements from collection management systems and documents. A formal requirements message from the U.S. Central Command had increased the requirement for GMTI to 10 times the current requirements. Thus, one of the team's initial analytics was focused on characterizing the detailed collection requirements. In doing so, the team learned that the collection requirements repositories had become laden with duplicative requirements over time. Since these collection management systems tend to have limited ability to analyze the total collection requirements, it is difficult to detect these redundancies without specialized analysis. The team analyzed all of the requirements

through several factors: spatial and temporal redundancy (roughly 30% of the requirements were duplicative), active requirements for events that had passed, and information need satisfaction by GMTI as three examples. This initial analytic winnowed such a large percentage of the requirements that it called into question the utility of the requirements repository.



In a similar analytic, a consultant deployed to Afghanistan examined the GMTI requirements of one Regional Command (RC). This RC had well over 1,000 requirements, but received very little collection. The consultant analyzed the requirements, leveraging ArcGIS and tailored processing techniques, against the terrain in the RC's area of operations and the stated information needs. In summary, this analytic identified a little over 100 requirements suitable for GMTI collection. With this focused and better-justified set of requirements, the RC was able to actually increase the amount of collection it received.

These GMTI assessments also illustrate the need to process a wide variety of data sets when conducting assessments. In addition to the larger volume of collection requirements, the team collected and processed End of Mission Summaries and thousands of target tracks and cross-cue events; raw GMTI collection files representing millions of dots; and hundreds of analytical products, produced from the GMTI dots, characterizing traffic patterns and pattern-of-life analyses. As was the case in the other assessments, these files represented a wide variety of data formats that required processing in order to render the data into a format suitable for analysis. As one example, the team gathered one year's worth of tracking data on U.S. forces for a special analytic. This required tailored code and ArcGIS processing distributed over several computers working in parallel to process the data. With this "blue force" and similar data, the team was able to analyze GMTI in the context of the Common Operating Picture.

One operational analytic characterized when GMTI provided cues on genuinely suspicious activities and when the cues were in effect false

alarms. To assess this issue, the team analyzed data derived from the JSTARS End of Mission Summaries which detailed every GMTI track created by on-board analysts, as well as cross-cue events (specifically, events where another ISR asset was cued to investigate a “suspicious” GMTI track). After parsing these reports, the team calculated the number of “suspicious” tracks generated. Then it isolated each cross cue incident to determine if the request for a cross-cue asset was satisfied and if the cross-cue asset confirmed that he reported activity was in fact suspicious. This analytic determined that out of thousands of GMTI tracks, cross-cueing of other sensors occurs about 20% of the time—a substantial number of times for a high-resolution sensor to be tasked to investigate a potentially suspicious activity. The analytic further determined that high-resolution sensors confirm suspicious activity in only 10% of cross-cued events. The other 90% were spurious cues.

Beyond the tremendous volume of empirical data, the assessment teams also collected a variety of information from direct observation and interviews in the war zones; a tailored survey; and documentation on the technique characteristics of WAMI technologies in development. While the assessment method emphasizes empirical data, the other data is a useful complement. For example, the tailored survey results highlight a disparity between intelligence collection managers—those that submit GMTI collection requirements—and the specialists in GMTI data. The GMTI specialists’ expectations of GMTI utility are much better aligned with the systems’ actual performance than with those of the collection managers.

In response to the JROCM, the team formed five primary analytic work streams—three focused on quantifying and improving current utility, and two focused on the potential utility of advanced processing and sensor capabilities. The latter two questions, emphasized by General Cartwright, yielded some of the most useful analytical results and reinforced an important lesson to these assessment teams.

Summary of the Results and Implications

The 2007 HVI assessment had demonstrated with hard data the severe limitations of GMTI’s value to irregular warfare. These additional analytics, occurring over a three-year period and addressing both war zones and a diverse mission set, only reinforce those limitations. Although the point has been made, it is worth noting that the GMTI was perceived as having high value to irregular warfare operations

based on a set of success-story briefings. These assessments dispel the success stories with hard data, but also identify a number of opportunities to increase the value of GMTI in response to the JROCM (Additional conclusions were derived for other WAMI sensors.).

GMTI sensors were found to have higher potential value to irregular warfare if the sensor resolution is increased substantially, even though this would require greater revisit time by the radar and therefore much less area collected. Some newer GMTI sensors are promising in this regard, while the traditional GMTI sensors remain ill-suited to irregular warfare needs.

By analyzing the value of multiple WAMI sensors and platforms in context of the earlier assessments of airborne FMV value, the team was able to construct a concept of operation for effective integrated cueing between GMTI and FMV sensors. This concept involves newer GMTI and FMV sensors with the GMTI sensing drawn closer to the platform to fully integrate with the FMV's field of view. The U.S. Army would eventually analyze this concept against the future irregular warfare needs in the Integrated Sensor Coverage Area (ISCA). The ISCA analysis would incorporate this integrated sensing approach as a key component of the Army's desired future architecture.

The team also identified an opportunity to automate a type of GMTI analytical product—traffic patterns—that consumes a large percentage of the GMTI analytical workforce. This product is relatively easy to automate and would free precious time for the analysts to concentrate on high-value tasks that cannot be automated.

The JROCM and interactions with General Cartwright reemphasize an important lesson for the assessment teams. While quantifying the value of intelligence capabilities with hard data is relatively new, it must not become the assessment team's end objective. Clients are not particularly well-served by detailed quantification of how poorly current systems are performing. They need solutions, or they must at least see analysis of alternative solutions.

About the Authors

Frank B. Strickland, Jr. is a Senior Fellow with the IBM Center for The Business of Government and a Partner in IBM's Global Business Services.

Prior to joining IBM, Mr. Strickland co-founded Edge Consulting, a consulting firm that achieved national recognition for pioneering work in the application of operations research methods and IT to quantify the value of intelligence. He helped lead Edge Consulting from a start-up to significant annual growth, culminating in its acquisition by National Interest Security Company.

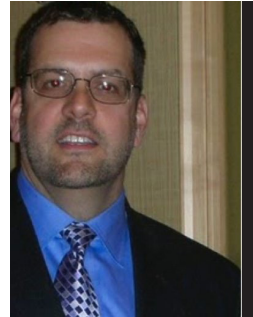


Mr. Strickland was a career intelligence officer with 24 years experience in the Central Intelligence Agency's Senior Intelligence Service and the U.S. Marine Corps, where he led programs focused on developing innovative solutions and methodologies to measure and analyze mission performance. In recognition of his accomplishments, the CIA Director awarded him with the National Intelligence Medal of Achievement. Mr. Strickland also received the National Reconnaissance Office's Medals of Distinguished and Superior Service.

Mr. Strickland is the co-creator of "Edge Methods," a unique blend of consulting, scientific methods, and IT used to assess the value of information from empirical data. Edge Methods has been used to advise national security principals and commanders on the optimal use of billions of dollars of operational and fiscal intelligence resources. He is a recognized teacher, public speaker, and published author. He holds a BA in Business Management, MS in Technology Management, and the CIO University's Certificate in Federal Executive Competencies.

Chris Whitlock is a Partner in IBM's Global Business Services.

Chris Whitlock has worked defense and national security issues for the past 30 years. For the last 20, he focused primarily on strategy consulting on intelligence issues from an analytic perspective. He co-founded and was the CEO of Edge Consulting, which applied empirical methods and management consulting techniques to advise on major programmatic issues confronting DoD and the Intelligence Community. He holds a B.A. in History (Mississippi), an M.A. in National Security (Georgetown) and an M.B.A. (George Mason).



Chris was trained as an Infantry officer in the United States Army and subsequently became a military analyst with the CIA. He is an expert in technical forms of sensing, having spent most of the past 20 years working projects to improve performance for imagery and signals intelligence systems. He has worked target problems in a variety of forms and countries, including Panama, Colombia, El Salvador, Bosnia, the former Soviet Union, Korea, Iraq, Iran, and Afghanistan. He has led and worked on projects cutting across a range of mission areas including counter-insurgency, theater missile defense, counter-fire, suppression of enemy air defense, combat search and rescue, and combined arms maneuver. With respect to methods, Chris has applied multi-attribute utility theory, modeling, and various empirical approaches including the development of a Warfighting Applications Research approach and what came to be known as "Edge Methods."



IBM Center for
The Business of Government

About the IBM Center for The Business of Government

Through research stipends and events, the IBM Center for The Business of Government stimulates research and facilitates discussion of new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

About IBM Global Business Services

With consultants and professional staff in more than 160 countries globally, IBM Global Business Services is the world's largest consulting services organization. IBM Global Business Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build, and run those solutions in a way that delivers bottom-line value. To learn more visit: ibm.com

For more information:

Jonathan D. Breul

Executive Director

IBM Center for The Business of Government

600 14th Street NW

Second Floor

Washington, DC 20005

202-551-9342

website: www.businessofgovernment.org

e-mail: businessofgovernment@us.ibm.com

**Stay connected with the
IBM Center on:**



or, send us your name and
e-mail to receive our newsletters.