Lessons Learned on

Information Technology Performance Management:

Applying the Balanced Scorecard

and

Applied Information Economics

to

Federal Information Technology Initiatives

Prepared by:

Federal Chief Information Officers Council

Best Practices Committee

Community of Practice for IT Performance Management

TABLE OF CONTENTS

Executive Summary

The U.S. government is more focused on results than at any other time. In 2001, the President established a management agenda to deliver greater results to the American public and place greater accountability on Federal executive departments and agencies. A key element of the president's management agenda is the use of information technology (IT) to achieve greater results and efficiencies.

Measuring the contribution of IT investments to complex missions such as providing for the health, welfare, and defense of the citizens of the United States requires Federal executive agencies to overcome significant technical and management challenges. These challenges are magnified by a lack of an established body of knowledge for determining IT's contribution and methods to measure it efficiently. Federal executive agencies also lack examples relevant to their missions to guide their measurement practices and help them overcome organizational resistance as they incorporate measurement into their management practices.

To address this knowledge gap, the Federal Chief Information Officers (CIO) Council completed experiments of two measurement methodologies applied to two major Federal IT initiatives in 2001. The two methodologies were the Balanced Scorecard (BSC) and Applied Information Economics (AIE). These pilot demonstrations showed that both methodologies determined and measured the contribution of specific IT investments to the missions of two Federal departments. They also showed that the methodologies are management tools for improving performance, though their focus and approach are quite different. Even though the pilots were completed more than two years ago, the performance measures developed would be still appropriate to include in annual performance plans and budget exhibits for capital investments. They also would be compliant with the Performance Reference Model developed by the Office of Management Budget in September 2003.

The BSC methodology clarifies the intended purpose of the IT investment, aligns it to the mission, vision, and objectives of the organization, and determines the strategy to implement it. The BSC metrics evaluate whether the strategy is being implemented and whether it is working. It did not provide tools to define measurement methods, however, to implement the metrics.

The AIE methodology improves investment decisions through a comprehensive analysis that quantifies costs, benefits, and risks in terms of dollars and probabilities of occurrence. AIE determines which costs and benefits to measure based upon whether additional information from measurements will improve a specific investment decision, estimates the costs of measurement and the degree of accuracy needed for each measure, and defines measurement methods.

The strength of the BSC methodology was that it aligned and fostered consensus for a proposed IT investment to the mission and strategic goals of a Federal department and three of its agencies. The BSC methodology defined strategies to achieve mission results and to manage the development and implementation of a food acquisition tracking and entitlement system, a system that supports food programs. The BSC pilot resulted in a Balanced Scorecard that included five perspectives with three objectives (performance goals) plus a total of 50 metrics and 39 initiatives to achieve the objectives. Not all metrics needed to be implemented initially.

The strength of the AIE methodology was that it defined measurement methods and quantified the costs, benefits, and risks for an information security program, considered difficult to measure. The AIE pilot developed and used seven performance metrics to identify the best combination of security program investments that would reduce the frequency and severity of IT security incidents that cause IT security-related losses. The AIE pilot identified a potential $30 million cost avoidance by determining that the benefits of one investment did not justify its cost. The pilot participants found the AIE methodology tough to comprehend due to its scientific and mathematical nature but they considered the performance measures highly useful.

The BSC methodology was stronger at determining and prioritizing organizational needs, defining a strategy to meet those needs, gaining buy-in, and linking initiatives to organizational strategy. The BSC methodology was also more intuitive. The AIE methodology was stronger at defining performance clearly to make it measurable, designing methods to collect measurement data, and providing guidelines to interpret the data.

Observers of both pilots believed that both methodologies provided value and that a combination of the methodologies would be ideal. Because of the cost of the methodology experts and the amount of involvement by government personnel, both methodologies are only cost effective for large IT investments. Implementing these methodologies requires an organizational commitment in resources and time that goes beyond the development of performance measures. To benefit from these methodologies, organizations need to commit to improving performance and adapting management practices to use performance information. Otherwise, these are expensive tools to merely report performance.

Steps Performed During the Pilots	AIE	BSC
Aligned IT Investment to Mission and Objectives of Agency	Yes	Yes
Prioritized Stakeholder and User (customer) Needs	No	Yes
Built Consensus within Agency	No	Yes
Quantified and Estimated Costs, Benefits and Environmental Factors	Yes	No
Determined Risk and Return of IT Investment	Yes	No
Developed Measures to Estimate Impact of IT Investment on Mission	Yes	Yes
Provided Methods to Maximize Impact of IT Investment	Yes	Yes
Prioritized Performance Measures	Yes	Yes
Designed Methods of Measurement	Yes	No
Identified Sources of Measurement Data	Yes	No
Designed Roll-Out Strategies to Maximize Return on Investment	Yes	Yes
Identified Initiatives for Acquisition, Development and Implementation of System	No	Yes

I. Background

The Government Performance and Results Act (1993) and the Clinger-Cohen Act (1996) require Federal executive agencies to measure the results of their major programs and to establish processes to measure the contribution of their information technology (IT) investments to mission results. In July 2000, the Capital Planning Committee of the Federal Chief Information Officers Council established the Subcommittee on Information Technology Performance Management to identify and address issues that would help Federal departments and agencies measure the contribution of information technology investments to mission results.

The Capital Planning Committee funded the subcommittee to conduct pilot demonstrations to determine if commercial measurement methodologies could measure the hbcontribution of IT in Federal agencies. The subcommittee chose the Balanced Scorecard (BSC) methodology because of its wide spread use in industry and growing use within the Federal government to measure organizational performance. The subcommittee chose the Applied Information Economics (AIE) methodology because of its claim to measure difficult to measure IT benefits such as improved decision making.

Completed in May 2001, the pilots showed that each methodology was applicable in the Federal environment. Shortly after publishing reports on the pilots in September 2001, the Federal CIO Council reorganized and established the Best Practices Committee to collect and disseminate best practices from the Federal e-government initiatives. In January 2002, the Best Practices Committee formed communities of practices for IT capital planning and performance management to facilitate the sharing of knowledge. Mr. Patrick Plunkett[1] from the Department of Housing and Urban Development observed both pilots and authored this report. The Best Practices Committee approved this report for publication.

II. Introduction

Although the pilots were completed more than two years ago, Federal agencies still face the same requirements and challenges to measure the performance of their IT investments. The purpose of this document is to provide federal agencies with examples that measure the contribution of IT investments to mission results. It also provides an in-depth look at each methodology and insights for developing and using performance measures effectively. This report begins with a summary of the major steps of each pilot followed by an examination of their similarities and differences along their strengths and weaknesses.

The Departments of Agriculture and Veterans Affairs volunteered to host the pilots plus the time of their project personnel, managers, and executives to participate. The USDA presented one of their major business line IT investments and the VA presented a major IT infrastructure investment. The following table identifies the IT investments used in each pilot, a brief description of each methodology and the results produced, and the amount of involvement by Federal personnel.

[1] Mr. Plunkett is also a co-chairman of the Community of Practice for IT Performance Management.

Summary of the Pilot Demonstrations

Host Agency	Dept. of Agriculture	Dept. of Veterans Affairs
Type of IT Investment	Replacement of business line system for three agencies (Food management system)	New department-wide infrastructure system (Security)
Status of IT Investment	Preparing for approval, not funded	Approved and minimally funded
Methodology	Balanced Scorecard (BSC)	Applied Information Economics (AIE)
Basis	Consensus among agency personnel on objectives, strategy, and measures	Mathematical and scientific principles and calculations
Focus	Define IT strategy to achieve organizational objectives	Analyze IT investment decision
Purpose of Performance Measures	Implement and validate strategy	Improve decision making
Approach	The contractor facilitated Government personnel to determine the strategy	Contractor completed an analysis based upon inputs from Government personnel
Knowledge and Skills Required by Government Personnel	Existing and proposed system, agency business processes, customer needs, agency priorities and needs, and MS PowerPoint	Security applications, processes and threats, impact on agency, and MS Excel
Results	A scorecard with five perspectives with three objectives per perspective, 50 measures, 39 initiatives, strategy map, and templates for defining measures and initiatives	Seven metrics and data collection methods, a $30M cost avoidance, post security incidence survey, rollout strategy for public key infrastructure, and recommendations for enhancing three optional investments
Duration	75 days	52 days
Average hours on project by Federal team members	92 hours (10 members)	66 hours (7 members)
Number of hours on project by Federal team leader	310 hours	350 hours

III. Overview of the Balanced Scorecard (BSC) Pilot

The Department of Agriculture (USDA) volunteered to participate in this pilot[2] with its Food Acquisition Tracking and Entitlement System (FATES) initiative. FATES was a proposed replacement system for the Processed Commodities Inventory Management System, which is used jointly by the Agricultural Marketing Service (AMS), Food and Nutrition Service (FNS) and Farm Service Agency (FSA) to administer the domestic and international food commodity programs funded at $1.5 billion annually. AMS, FNS, and FSA are referred to as the Tri-agencies.

> "A good scorecard is more than a collection of metrics. It tells the organization's strategy."
>
> Debra Whitaker
> BSC for Government, Inc.

The BSC is a top-down methodology that examines organizations from internal and external, financial and non-financial, and short- and long-term perspectives. The philosophy of the BSC is that organizations are more effective when guided and aligned by their mission and vision and when focused on multiple perspectives. A good balanced scorecard is a mirror of an organization's strategy. The performance measures translate the strategy into action. The term "balance" comes from an examination of multiple perspectives instead of a single financial perspective. It is not "balanced" in a mathematical sense where perspectives are assigned weights to calculate a final score.

The contractor, The Balanced Scorecard For Government, Inc., used a tailored version of the Balanced Scorecard methodology created by Harvard Professors Dr. Robert Kaplan and Dr. David Norton. The tailored Balanced Scorecard for government agencies consisted of the following five perspectives:

- Customer Perspective
- Stakeholder Perspective
- Internal Perspective
- Learning and Growth Perspective
- Financial or Budgetary Perspective

Based upon the contractor's recommendations, each perspective included three objectives with measures to gauge progress towards each objective, plus initiatives for achieving the objectives.

Preparing for Development of the Scorecard

The pilot began with the contractor advising the USDA to assemble a team that included senior IT and program specialists. The USDA selected ten IT and program specialists responsible for the existing system as well as for the development and implementation of the FATES. The contractor also advised the USDA to form an executive team to decide the contents of the

[2] For more information, see Measuring IT's Contribution to Mission Results: A Case Study of the Balanced Scorecard Methodology for a Business Line IT Investment (http://cio.gov/documents/bsc_report_final.pdf)

FATES BSC. For the executive team, the USDA chose the senior executives from the Tri-agencies who were members of the advisory committee for the existing food commodity system. During this initial stage, the contractor reviewed strategic plans, annual performance plans and documentation on the FATES provided by the USDA.

Defining the Strategic Objectives

To begin the development of the FATES BSC, the USDA Team created a mission and vision statement for FATES that linked to the mission statements of each Tri-agency and to the USDA. The USDA Team used these statements to develop a list of objectives for the Customer and Stakeholder perspectives. Before developing the objectives, the contractor helped the USDA Team to identify and reach consensus on the major internal and external customer segments and stakeholders of FATES would be. To create the list of objectives, the USDA Team answered this question: "What must FATES accomplish for its customers and stakeholders that will help achieve the mission and vision for FATES?" The contractor used templates with examples to stimulate brain storming among the USDA Team.

Next, the USDA Team defined objectives for the Internal perspective by answering the question: "What must the USDA Team do internally to achieve the objectives of the Customer and Stakeholder perspectives?" For the Learning and Growth perspective, the USDA Team answered the question: "What people, skills, tools, and culture do end users need to use FATES and to achieve the objectives of the Internal perspective?" For the Financial Perspective, the USDA Team answered the question: "How is FATES funded? And how will the funds be managed?" Members of the USDA Team suggested objectives and reached consensus on the list of objectives for each perspective.

To determine the top three objectives for each perspective, the contractor created an interview guide that included the mission and vision statements plus the list of objectives for each of the five perspectives. The contractor along with members of the USDA Team used the interview guide to conduct structured interviews of senior executives and managers within the Tri-agencies who had an interest in FATES and the food commodities program. There were 23 separate interviews. Each person interviewed evaluated the FATES mission and vision statements, ranked the top three and sometimes four objectives for each perspective, defined the characteristics of good performance for each objective, and suggested ways to measure the objectives.

To determine which objectives that represented the consensus among of those interviewed, the contractor calculated a weighted score for each objective within each perspective. This involved multiplying a weighting factor (e.g., a '1' equals 10 points, a '2' equals 8 points, a '3' equals five points, a '4' equals 2 points, not chosen equals 0) by the number of times selected. For example, if the objective "Timeliness" received a rank of '1' by 2 interviewees, a '2' by two more interviewees and a '3' by one interviewee, its weighted score would be 41 (i.e. 2 times 10 plus 2 times 8 plus 1 times 5). For each perspective, the three objectives with the highest score represented the consensus of the executives and managers. The USDA Team presented the results of the interviews to the Executive Team, which could either reject or approve the

rankings. After receiving approval for the objectives, the USDA Team created a formal definition for each objective.

Developing Measures and Initiatives for Each Strategic Objective

After the objectives were defined, the contractor provided a tutorial for developing performance measures. The USDA Team then separated into small groups and formulated possible measures for each perspective. The USDA Team recorded each measure onto a measure template (Appendix A). The team reconvened and examined each measure to determine how useful the information from the measurements would be to evaluate progress towards the objectives.

> "A Balanced Scorecard without initiatives is a worthless scorecard. Without clearly defined initiatives, the objectives are meaningless."
>
> Debrah Whitaker
> BSC For Government, Inc.

After reaching consensus on the definitions for the measures, the USDA Team identified initiatives necessary for collecting the measurement data and for achieving the objectives. The USDA Team recorded the initiatives using measurement templates (Appendix B). After the USDA Team presented the measures and the initiatives, the Executive Team approved the scorecard with the understanding that further refinement of the metrics would be done. See Appendix C for the five perspectives of the FATES Balanced Scorecard.

Implementing the Scorecard

According to the contractor, the key to implementation was defining the purpose and formula for all measures, specifying data sources, and assigning roles and responsibilities for setting targets, meeting the targets, and tracking and reporting the results. The contractor advised that not all measures should be implemented at once. During the first year of implementation, the contractor recommended that the USDA Team refine and prioritize its measures, set up data collection processes, and establish baselines for the highest priority objectives. The contractor recommended that the USDA Team focus on implementing the high impact objectives first because of the cost of measurement and overall constraints on resources. The contractor also advised the USDA Team, working with the FATES Advisory Council,[3] to assign ownership of the balanced scorecard and to encourage that the roles be carried out.

The contractor said ownership consisted of three roles. For each perspective, a champion was needed with the authority to change the objectives and measures and with the authority to task people with initiatives. The champion's responsibility was to ensure that the data for the measures are collected timely and correctly. Each owner of a row of scorecard, that included the objective, lead and lag measures, targets, and initiatives, was responsible to collect and report the data to the champion. The owners of initiatives were responsible for the initiative's completion. Each role was important to achieving results and for keeping the FATES BSC useful.

[3] Composed of one executive from each of the Tri-agencies.

To avoid increasing workload and creating a barrier to implementation by adding the BSC responsibilities to existing management practices, the contractor recommended that the USDA team integrate the scorecard into existing management practices within the Tri-agencies. To accomplish this, the contractor recommended that the USDA team identify existing management practices that might no longer be needed. To gain acceptance, the contractor advised the USDA Team to actively communicate the content of the scorecard to managers and the persons who work on FATES and the reasons the BSC was established. The USDA Team will also increase acceptance of the scorecard by soliciting input from those who work on FATES on the content of the scorecard.

In the second year of implementation, the contractor advised the USDA Team to focus on using the BSC to manage FATES and complete the implementation of the scorecard. According to the contractor, the better ways to do this included aligning the BSC to the budget and using the BSC as a filter to judge which sub-projects to fund. The contractor recommended that the USDA Team update the scorecard and work on overcoming barriers to implementation. The contractor warned that if the USDA Team did not use the BSC in its management practices to make decisions, then implementation of the FATES BSC would stall because this would signal to those responsible for collecting and using the measurement data that the scorecard was no longer important. As the Tri-agencies collect measurement data, the contractor indicated that the BSC could serve as a tool for the USDA Team to examine the strategies for FATES, determine if they were being successfully implemented and if the strategies were working. The contractor advised the USDA Team to make changes to the BSC when strategies changed.

IV. Observations and Lessons Learned from the BSC Pilot

Organizations that have balanced scorecards use them primarily for their business units. A BSC for an organization's IT function or a major IT initiative such as FATES would normally be developed after the organizational scorecards were in place to ensure that they were aligned with organizational objectives. Because no organizational BSC existed for the Tri-agencies, the USDA Team had to clarify organizational objectives in addition to the IT objectives. The contractor stated that BSCs are more effective when organizations cascade them down and throughout organizations to set priorities and guide objectives of operating units.

Differences Between Government and Industry Balanced Scorecards

The FATES BSC contained a fifth perspective, "Stakeholder." This differed from the BSCs developed by Kaplan and Norton for private industry firms, which typically have four perspectives: Financial, Customer, Internal, and Learning and Growth. The "bottom line" for the private industry is the financial perspective. For the FATES BSC, the contractor included the Stakeholder perspective to address the mission and legislative mandates of USDA, essentially its bottom line. The Stakeholder Perspective, which reflected the interest and priorities of stakeholders, focused on program delivery, operational efficiency, and effectiveness of program administration.

The architecture of the government Balanced Scorecard was also different. The mission and vision statements for FATES "drove" the objectives of the Stakeholder and Customer perspectives. The objectives of the Stakeholder and Customer perspectives drove the objectives for the Internal Perspective, which drove the objectives of the Learning and Growth Perspective. The objectives of the Financial perspective focused on cost savings in food products and services and better use of funds for recipients of the food programs. With commercial BSCs, the mission and vision statements drive the objectives of the Financial perspective that drive the objectives of the Customer perspective that drive the objectives of the Internal perspective that drive the objectives of the Learning and Growth perspective.

Differences Between IT and Organizational Scorecards

When developing the FATES BSC, the USDA Team examined each perspective from both an internal and external viewpoint. The USDA Team identified internal and external stakeholders and customers. For the Customer perspective, the USDA Team selected 'Better Information' as an objective because the existing system did not provide sufficient or timely information for internal USDA users who serve program recipients. By providing better information to internal users, the USDA Team believed that the Tri-agencies would provide better service to both its program participants and recipients. The USDA Team chose 'Customer satisfied with level of service" as an objective for the Customer perspective because of the importance of external customers. For organizational scorecards, only the Customer perspective examines the external viewpoint.

Distinctions Between Objectives in the Customer and Internal Perspectives

In the course of developing possible objectives, the USDA Team identified the objective "Flexible System" for the Customer Perspective and the objective "Flexibility to Implement Change in a Timely and Cost Efficient Manner" for the Internal Perspective. Initially, it was unclear whether these objectives were duplicative. The contractor clarified the matter by explaining that each of these objectives represented a different "window" into the FATES. The objective "Flexible System" represented the customers' view, i.e. what they wanted the FATES to do. The objective "Flexibility to…" represented the Tri-agencies' internal view, i.e. the flexibility of internal processes to meet the needs of specific customers.

The Importance of Consensus

Consensus is a fundamental principle of the BSC methodology. Its inclusion in the BSC evolved from the early experiences of Kaplan and Norton. Within six to twelve months after Kaplan and Norton had developed the scorecards for their clients, many stopped using their scorecards. After conducting research and consulting with experts in change management, they incorporated consensus building into their methodology. Although reaching consensus took considerably more time, their clients took more ownership of the scorecards, implementation improved, and use continued for years. Kaplan and Norton believed that consensus or "buy in" during development of the BSC was key to adoption and successful implementation.

During the pilot, members of the USDA Team reached consensus on the mission and vision statements, objectives for each perspective, performance metrics, and on the initiatives to achieve the objectives. Although the team members represented three agencies and included both program and IT specialists, the team reached consensus fairly easily since they had worked together on a regular basis since the early 1990's. For organizations that have not worked together as long, consensus would likely take more time to achieve because they would not have the confidence and trust in one another to collaborate on important matters of strategy and performance measurement.

Members of the USDA Team expressed surprise at the high degree of consensus among their executives and managers, particularly since more managers from one agency were interviewed than from the other two. Initially, some members of the USDA Team thought this imbalance would skew the results. To the contrary, the results of 23 interviews revealed that the managers and executives shared essentially the same opinions regarding the top three objectives for each perspective.

The Interview Process

The interview process was the principal method for determining the objectives for the FATES BSC. The contractor and members of the USDA Team conducted structured interviews of 23 managers and executives from the Tri-Agencies. The USDA Team selected these managers and executives on the basis of their responsibilities and interest in FATES.

The USDA Team identified three customer segments, Internal (USDA Users), External (Non-USDA Users), and External (Direct Recipients) prior to developing a list of objectives for the Customer perspective. The USDA Team considered interviewing external customers but decided that the managers who were internal customers were sufficient because of their knowledge of the needs of all customers. The contractor advised that choosing candidates to interview always involved tradeoffs between the time and cost of conducting interviews and the value of additional information.

Lead and Lag Indicators

The FATES BSC included two types of performance measures or indicators. The USDA Team developed a lag measure for each objective to determine the degree it would be accomplished over time. A lead measure was developed when a significant change in strategy or performance was needed. These measures have a cause-and-effect relationship. For example, in the Customer perspective for the objective "Better Information," the lag measure developed was "Reliability of Data in FATES." The lead indicator was "Decrease in requests to IT staff for reports." The USDA Team believed that a decrease in the number of requests would signal that users were getting better information. Not all objectives needed a lead indicator but all objectives needed a lag measure. The difference between the indicators and how to use them was not obvious. During implementation, the cause-and-effect relationships need to be validated because lead indicators prioritize resource allocations.

V. Overview of the Applied Information Economics (AIE) Pilot

The Department of Veterans Affairs (VA) participated in this pilot[4] with its Information Security Program (ISP), an approved new infrastructure initiative intended to enhance information security across the Department. In January 2001, the VA estimated that the total cost for ISP at approximately $114 million. For this pilot, the VA provided a team that consisted of a team leader and six senior IT-security specialists.[5]

Applied Information Economics, created by Doug Hubbard, President of Hubbard Decision Research, Inc., is a quantitative decision-analysis tool that uses a synthesis of techniques from the fields of economics, financial theory, decision theory,[6] information theory,[7] and statistics. The AIE has multiple uses.[8] For this pilot, the contractor used AIE to determine the performance measures that will provide information to help the VA make the best investment decision on the ISP. The same measures also will determine the ISP's effectiveness.

The Clarification Process

The contractor began the pilot with a review of the VA's strategic plan, annual performance plan, IT Strategic Plan and documentation on the ISP to identify specific objectives that the ISP would contribute. The contractor used this information to clarify the purpose of the ISP and its major investments or components[9] and later to define the investment model. The contractor clarified the purpose of the ISP by working with the VA Team to uncover what was meant by enhance information security. After examining the purpose of each investment, the contractor formulated that the purpose or the value of the ISP is to reduce the frequency and severity of three types of security incidents (viruses, unauthorized intrusions, and environmental events) that cause four types of losses (fraud, productivity, interference with mission, and legal liability). The diagram below illustrates the purpose or value for the ISP.

Translating this into the language of performance measurement, the outputs of the ISP were a reduction in the frequency and severity of IT security-related incidents. The outcome of the ISP was a reduction in losses from IT security–related incidents.

The contractor also found that each ISP investment consisted of a minimal and an optional component. Since the VA had decided to fund the minimal components of each ISP investment, the contractor concluded that the investment decision to analyze was which combination of optional components would reduce the most security-related losses.

---

[4] For more information, see Measuring IT's Contribution to Mission Results: A Case Study of the Applied Information Economics Methodology for an Infrastructure IT Investment (http://cio.gov/documents/aie_report_final.pdf)

[5] Seven people participated part-time for a total of 300 hours over a five-month period.

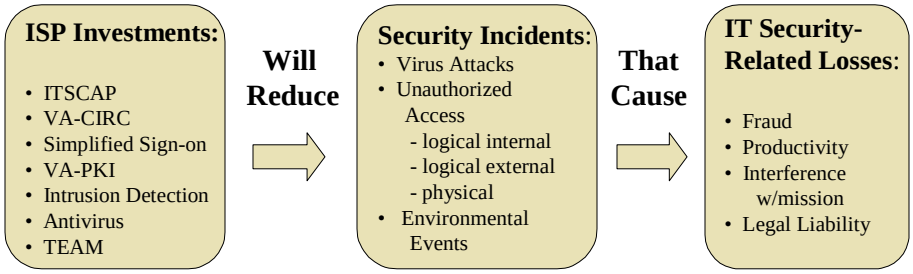
[6] The formulation of decisions into a mathematical framework

[7] The mathematical modeling of transmitting and receiving information.

[8] AIE is also used for portfolio management and risk-return analysis.

[9] The seven ISP investments were Public Key Infrastructure, Intrusion Detection, IT Systems Certification and Accreditation Program, Simplified Sign-On, Antivirus, Computer Incident Response Capability, Training, Education, Awareness, and Message Building. VA completed a Threat Assessment investment prior to the pilot.

Value of VA Information Security Program



Legend:

ITSCAP: IT Systems Certification and Accreditation Program
VA-CIRC: VA Computer Incident Response Capability
VA-PKI: VA Public Key Infrastructure
TEAM: Training, Education, Awareness, and Message Building

The Investment Analysis

Before determining the best combination of optional investments to reduce the most security losses, the contractor first estimated the amount of IT security related losses that would likely occur if the ISP did not exist. The contractor worked with the VA Team to create a model using a Microsoft Excel spreadsheet that represented the security threats over a six-year period. Using calibrated estimates[10] provided by the VA Team for the frequency and severity of incidents, the contractor estimated the total losses without the ISP at $1.1 billion to $2.5 billion. The analysis also revealed that 71 percent of the security losses would likely come from unauthorized internal access compared to only 17 percent of security losses due to losses in productivity from a computer virus attack. This indicated that intrusion detection software would reduce more losses than anti-virus software.

> "The investment decision that VA faced was a classic maximization problem, i.e., which combination of ISP investments will reduce the most security losses."
>
> Doug Hubbard
> Hubbard Research, Inc.

The contractor then created an investment model to estimate the reduction of losses from each ISP investment and from different combinations of the investments. The contractor's estimation approach was innovative. Instead of using a single estimate for each variable, as is commonly

[10] The contractor calibrated the VA personnel to estimate their uncertainty so that they could provide a 90 percent confidence interval, i.e. a lower and upper bound, for a quantity that represented their 90 percent confidence level.

done for IT investment analysis, the VA Team, with facilitation and training in estimating their uncertainty, defined a range including a lower bound, a formula or best estimate, and a higher bound for each variable. The team members were 90 percent certain[11] that the actual value of the variables would fall within those ranges. For example, for the Productivity Impact of a Virus Attack variable, the productivity loss per incident ranged from 15 percent to 60 percent with 38 percent as the best estimate. The actual productivity loss per incident may be within this range or outside. The range estimates are more accurate than a single estimate because the probability of getting a precise number is zero.

To improve their estimates, the contractor taught the VA Team to factor their uncertainty[12] into the range estimates. To reduce their uncertainty, the contractor taught the VA team to conduct secondary research of the Internet and the Library of Congress for existing reports or studies that might contain findings that would help the team set the lower or upper bounds of their range estimates. For example, the VA needed to conduct secondary research because the VA's Inspector General annual report on fraud losses did not separately identify losses due to IT security-related incidents. The team found a report on fraud, waste and abuse in the Federal government issued by the Senate Governmental Affairs Committee that helped the team to significantly adjust the estimate for the upper bound for annual losses within the VA.

The contractor also taught the VA Team to select one of five possible probability distributions[13] for each variable. The contractor used the ranges plus the probability distributions with a Monte Carlo simulator to generate 50,000 random scenarios for cost and benefits, each a potential ISP outcome. The simulator kept track of the scenarios and plotted a histogram of the returns on investment along with the risk or uncertainty associated for all outcomes generated.

To prevent double counting of the benefits, the contractor modeled the combined effect of the investments. For example, both the Intrusion Detection and VA-PKI investments will detect and prevent unauthorized intrusions. If hypothetically, Intrusion Detection prevented 50 percent of the intrusions and PKI prevented 80 percent, their combined effect could not exceed a 100 percent reduction. If the effects of these combinations were added without adjustment, the model would calculate more than 100 percent of intrusions, which is not possible.

The simulator estimated the total losses from the best combination of ISP investments to be $100 million to $500 million--a reduction of 75 to 95 percent of the losses that would likely occur without the ISP. The analysis also revealed that the optional investment in Intrusion Detection did not reduce enough losses to justify its cost. This represented a potential $30 million cost avoidance for the VA. In addition, the contractor defined a rollout strategy for VA's Public Key

[11]The contractor improved and recorded the team's ability to estimate their uncertainty through training exercises.
[12] The greater the uncertainty, the wider the range between the lower and upper bounds or estimates will be. Interestingly, the contractor said that in his experience managers are overly optimistic when specifying ranges than workers are. The manager's ranges tend to be too narrow which increases risks.
[13] The probability distributions indicate the likelihood of the values in and outside the range. For example, a uniform distribution indicated that all values had the same chance of occurring. A normal distribution or bell curve indicated that an average value and values close to the average would likely occur most often. When a random number is generated for each variable, i.e. a random scenario, the selected probability distributions determined the values for each variable.

Infrastructure investment that will optimize its benefits by prioritizing the installation of PKI in the parts of VA that have higher annual losses.

Determining Variables to Measure and How to Measure Them

Using the investment model, the contractor determined which variables to measure by calculating the dollar value that additional information for each variable would have on reducing the uncertainty on the amount of security losses prevented by the ISP investments. By reducing the uncertainty about the amount of losses reduced by each optional investment, the VA would be better able to choose the best combination of optional investments that would reduce the most losses.

The contractor calculated the value of information for each of the 105 variables of the model using the Expected Value of Perfect Information (EVPI) [14] method. The value of information or EVPI for each variable represented the maximum dollar value that additional information would provide even if that information were perfect and there was no uncertainty about a particular variable. For example, if the VA knew the exact amount (perfect information) of productivity loss from virus attacks, they could determine which anti-virus investment would be most cost effective.

The variables within the model with significant[15] EVPIs were the ones worth measuring because the measurements would provide information that would significantly reduce the uncertainty about the investment decision. They included fraud, the costs of the optional investments, the frequency of logical intrusions, and the total effect of viruses and the reduction of viruses due to the ISP investments. These may appear somewhat obvious but according to the AIE expert, it is not intuitive before running the EVPI calculations which variables will have the higher information values.

Key Metrics for Improving ISP Investment Decision

- Fraud losses per year
- Intrusions per year
- Pandemic viruses per year
- Number of VA personnel affected per virus outbreak
- Duration of productivity impact per virus outbreak
- Average annual cost per person affected
- Productivity loss during duration

The EVPI calculations revealed that not all measurements were equal in value because the variables had different information values. Most variables had little to no information value meaning that if additional information was collected, it would not improve the investment decision. For the variables with significant information values, the contractor defined the seven metrics shown in the box. Not only were the metrics used to improved the investment decision, the VA will use these metrics to judge the effectiveness of the ISP to reduce the frequency and severity of IT related security incidents and the losses they create.

[14] From the field of information theory, according to Doug Hubbard, this method has existed for more than 50 years.
[15] The significant information values ranged from $151,910 for the total effect of viruses and reduction of viruses to ISP investments to $787,763 for fraud, property loss, and legal liability. All other variables had information values less than $1,000. Many had information values of $0.

The contractor also used the EVPI to determine the maximum amount to spend on each measurement. The contractor recommended that the VA should spend no more than 2 percent to 20 percent of a variable's EVPI on measurement annually based upon his past experience that indicated there are diminishing return spending more on measurement.[16] The actual recommended amount was closer to 2 percent. The contractor designed the measurement methods based upon the amount of uncertainty they would reduce plus the cost of the measurements so that the measurements would be cost effective. According to the contractor, there is a direct relationship between the accuracy of measurements and the cost of the measurements: the higher the accuracy the higher the measurement costs will be. The higher the accuracy the less uncertainty will exist. The amount of uncertainty that is acceptable depends on the decision to be made. See Appendix D for a complete description of each measure and the measurement methods.

For some of the measurements, the software would collect the measurements. For example, the Intrusion Detection software would produce a report of the number of unauthorized accesses. This would indicate the effectiveness of the PKI component when the VA compared the number of intrusions detected in the parts of VA that deployed the software to other parts that did not. The anti-virus software could generate a report on the frequency of viruses that would indicate the effectiveness of the TEAM and VA-CIRC investments in the same manner.

VI. Observations and Lessons Learned from the AIE Pilot

The philosophy and principles behind the AIE methodology challenge the conventional wisdom of performance measurement and require paradigm shifts in thinking to comprehend it.

A New Measurement Philosophy

The contractor believed that the purpose of measurement is to collect information to improve future decisions not to justify past decisions. The only exception was mandatory reporting or fulfilling oversight requirements.[17] At first, this philosophy seems to differ with the Government Performance and Results Act (GPRA) measurement paradigm that measures are used to determine the effectiveness of IT investments and Federal programs. Although the measures developed during the pilot were designed to improve the ISP investment decision, the measures also will determine the effectiveness of the ISP over time by comparing the measurements taken to improve the investment decision with the measurements after the investment is operational.

A New Measurement Approach

The contractor stated that measurement consists of three elements: concept, object, and methods. Understanding the distinctions between these elements is key to understanding and creating useful performance measures. The contractor's methodology is based upon the scientific

[16] The contractor explained that increasing spending on measurements above these percentages does not reduce uncertainty enough to warrant the additional expense.
[17] For the Federal government, executive departments and agencies are required to collect information about past decisions to allow for adequate oversight.

concept of measurement, i.e. "to reduce uncertainty about a quantity through observation." The contractor stated that when most people think of measurement, they think in terms of determining an exact number. Many decisions can be made without knowing an exact number. For example, do most people need to know the exact price of a Rolls Royce Silver Shadow to know whether they could afford it or is the knowledge that it costs more than $100,000 or even $75,000 sufficient? Because obtaining an exact number is almost always very difficult and expensive to do, most people conclude that many things, such as "improved decision making," are impossible to measure.

According to the contractor, another reason that so many objects appear to be "impossible" to measure is because they have not been adequately defined. AIE is based upon the assumption that if a benefit or a cost is defined unambiguously, then it is measurable. The contractor used the following rationale to approach measurement:

> If it is "better," then it is different in some relevant way.
> If it is relevantly different, then it is detectable.
> If it is detectable, then it is detectable in some amount.
> If it is detectable in some amount, then it is measurable.

The contractor used this clarification method to define the benefits of the ISP and did not consider the performance measures until all the variables were defined. More than half of the duration of the pilot was spent identifying what to measure, i.e. the specific results desired. Through the contractor's facilitation, the VA Team determined that increased security meant reducing the frequency and severity of IT security-related incidents. In response to the question "What is the benefit of reducing the frequency and severity of security incidents?" the VA Team concluded the benefits would be fewer losses. By defining losses further they became simpler to measure.

To estimate the productivity impact of a virus outbreak, the contractor recommended that a structured telephone survey to be taken immediately after a virus attack. To determine the number of viruses received, the contractor recommended using the anti-virus software. Other methods recommended included using annual fraud loss figures from reports published by the VA's Inspector General. The contractor designed and prioritized the measurements drawing upon his knowledge and experience.

> "Not all measurements are equal. The higher the accuracy, the higher the measurement costs will be. The higher the accuracy, the less uncertainty will exist. The amount of certainty should determine the level of measurement accuracy needed for a decision."
>
> Doug Hubbard
> Hubbard Decision Research, Inc.

New Quantitative Methods: New Paradigms

The use of range estimates and the quantification of uncertainty challenged the VA Team members' concepts of estimation. Much of the mental challenge seemed to come from not understanding how the range estimates could be used in the calculations. In other words, how

can the entire range of all 105 variables be combined in the calculations? The Excel software did not actually use the entire range of values in a single calculation but rather generated a random number for each variable (according to one of the five probability distributions selected) and kept track of the result. The result represented one possible outcome. Using a Monte Carlo simulator, this was done 50,000 times. The software plotted the results of the calculations showing the frequency of the results that occurred. The frequency indicated the probability that each outcome would likely occur.

Since using a single estimate for costs and benefits is a common and accepted business practice, it was not intuitive that using single-point estimates might produce bad results. But in quantitative analysis, the use of ranges is recognized as far superior to "point" estimates because they factor uncertainty, which always exists. For the ISP, the VA Team was uncertain of the exact cost of each optional investment would be over a five-year period, and of the number and severity of security related incidents that would occur or of the amount of losses that would be accrued. Given this uncertainty, if the VA Team had used single point estimates for benefits those estimates would almost certainly be wrong.

To illustrate this concept, consider if you were faced with the decision of whether to invest $1,000,000 on a computer system. Assume that a cost and benefit analysis using single point estimates determined that the return on investment to be 50 percent over cost. You may conclude that the investment was worth making. If on the other hand, a cost and benefit analysis using ranges and probability distributions determined that there was only a 60 percent chance of a 50 percent return and that there was a 30 percent chance of a negative return, is the investment worth making? The answer would depend on your tolerance to risks and other investment alternatives available to you. The range estimates provide more information regarding risks to make a more informed decision.

At this point, it may be unclear what improved estimation and analysis of investment decisions have to do with measuring the contribution of an IT investment. By modeling the investment decision using range estimates for the costs and benefits (contribution), more accurate range estimates for total costs and losses were obtained. From this information, the contractor determined which combination of ISP investments would reduce the most losses considering costs and risks. Once this information was determined, the contractor calculated the impact of specific measurements would have to improve the investment decision by reducing the uncertainty or the width of the range estimates for losses. The variables or objects that reduced the greater uncertainty were the ones that justified the cost of measurement. It is interesting to note that additional information (additional measurements) on only 7 variables out of 105 would reduce the amount of uncertainty of total possible losses.

Principles of AIE

Information reduces uncertainty. Less uncertainty improves decisions. Better decisions result in more effective actions. And more effective actions improve mission results.

It took the VA Team and observers from the Federal CIO Council time and much discussion with the contractor to come to terms with this concept for selecting variables to measure using

the value of information. This was due in part because in information security, the value of information refers to the cost of protecting and replacing information that has been destroyed or changed. It was also easy to get lost in the mechanics of building the model. But understanding the mechanics of the AIE methodology was not necessary to understand the metrics defined. It was clear that someone that is proficient with spreadsheet software is needed to build the model.

Measurement Discoveries

Prior to the pilot, the VA Team conducted research to find performance measures for information security. The recommended measures they found were "the number of workers that completed security training" and "the number of times senior managers were briefed on IT security." Not surprisingly, the conventional wisdom among IT security professionals was that "enhanced security" could not be measured directly. The measures developed by the contractor contradicted that view. It was clear that the variables and their respective measures identified by the pilot would not have been discovered by the VA on its own. The VA Team was so impressed by the measures developed that they believed that other agencies could use them to meet the requirements of the Government Information Security Reporting Act.

VII. An Examination of the Pilot Methodologies

Before the pilots began, it was unclear whether the Balanced Scorecard and Applied Information Economics methodologies would be viable for measuring the contribution of IT initiatives to the mission of Federal agencies. The pilots demonstrated clearly that they were viable and could be useful to Federal agencies. The pilots enabled the Federal CIO Council to learn more about performance measurement by examining their similarities and differences as well as the their strengths and weaknesses.

Similarities between the Methodologies

Each methodology required the host agency to commit a small team part-time for four months to complete the pilots. The USDA and VA Team members alternated their time working on the pilots and their main duties. Because of the resource and time commitments involved with these methodologies, it is clear that organizations must be motivated to make this commitment. Both host agencies were highly motivated. The VA believed that they could improve the management of their IT Security Program and receive endorsement from oversight organizations with sound performance measures. The FATES program managers needed to establish a business case and performance measures in order to obtain approval and funding to proceed to the acquisition phase.

Each pilot began with a review of the agency strategic and annual performance plans and program documentation. Each methodology examined the organization's objectives to clarify the purpose and benefits of the IT investment. This determined the variables or attributes useful to measure. Both methodologies share the philosophy that everything is measurable. The clarification and definition steps, which accounted for more than half the time of each pilot, were

completed before performance measures were considered. Each methodology served as a lens to view and examine the IT investments and their contribution to mission results.

Using the AIE method, "increased security" was defined for the VA as "a reduction in the frequency and severity in the IT security incidents that cause losses in the form of fraud and productivity losses." The BSC method defined that "an efficient and effective FATES" for the USDA would mean better delivery and an effectively administered food commodity programs. It also would mean better information, a more flexible system, and better service for program participants and recipients.

Both methodologies developed a set of "critical-few" measures for making decisions and meeting oversight responsibilities. The AIE chose performance measures mathematically, based upon a trade-off between costs and the level of accuracy needed. The BSC chose performance measures for the top three objectives of each perspective through consensus among the members of the USDA Team and executives of the Tri-agencies. Furthermore, the BSC contractor recommended that the USDA measure the high-impact objectives first to leverage resources. Each methodology focused on the implementation and use of the performance measures to improve performance.

Differences between the Methodologies

Although the AIE and BSC are similar in some respects, their methods are quite different. The BSC focused on strategy definition and execution whereas the AIE focused on defining and analyzing key investment decisions. The AIE contractor determined which variables to measure, such as "the number of viruses and unauthorized intrusions and the losses that result," through a mathematical analysis. From a performance measurement perspective, the most significant difference was the AIE contractor prescribed specific methods for collecting measurement data but the BSC contractor did not.

The BSC contractor facilitated the USDA team in determining the objectives to measure such as "establish a good order entry system and provide users with better information" by a process that obtained consensus among members of the USDA Team and the Executive Team. The USDA Team defined performance measures for the objectives using guidelines provided by the BSC contractor. The USDA team used their judgment to select lead and lag measures based upon the measures that seemed to be "a good fit" with the objectives. The BSC contractor provided a template to define the formula, data source, and other information necessary for implementation.

The AIE contractor developed seven metrics for determining the best combination of optional investments to fund and for evaluating the effectiveness of each investment after implementation. With the BSC contractor's assistance, the USDA Team defined 15 objectives, developed 50 metrics, and identified 39 initiatives/tasks to implement the FATES strategy and to determine its effectiveness after implementation. The USDA Team planned to refine and reduce the number of metrics.

Strengths and Weaknesses of the BSC

A balanced scorecard and the process to develop one are somewhat intuitive. People readily accept the concept that multiple perspectives provide a more complete view of an organization's or IT initiative's performance even though the cause-and-effect relationships have not been verified. Including the Stakeholder perspective that addressed mission requirements made the BSC a better fit for government organizations than the standard four perspectives used in the private sector.

The strength of the BSC methodology is the manner in which it aligns IT investments to mission objectives. It also defines objectives for each perspective, links the perspectives to form a results-focused strategy, and builds consensus for the objectives and performance measures among not only the development team but also among key stakeholders. Consensus increases the likelihood that the performance measures reflect true priorities and that they will be implemented and used. The BSC also is a useful tool to communicate strategy and priorities throughout the organization.

A good BSC determines whether the strategy of an IT initiative is working or if it needs to be changed. The BSC provided a framework to define the results needed, i.e. FATES objectives, and to manage progress towards the objectives. It lacked rigor in places, however, particularly for determining the performance measures. The USDA Team chose the measures based upon the members intuition, judgment, and knowledge of the business process of the Tri-agencies. The BSC methodology provided no means by which to determine the correlation between the lead and lag measures and between the measures and the objectives. These cause-and-effect relations are assumptions that need to be verified. This lack of rigor could produce performance measures that are not useful and may cause organizations to spend more on measurement than is necessary.

Strengths and Weaknesses of the AIE

The strengths of the AIE included a process that clarified the ambiguous benefits of increased security, techniques that quantified uncertain costs and even more uncertain benefits, and methods that determined the usefulness of each performance measure. One of its unique strengths was its ability to design measurements that factored in cost and utility. It also estimated risk in a meaningful way, included it in the decision, and identified means to reduce it.

The mathematical rigor of AIE and its innovative approach to measurement were challenging to grasp and will likely impede its acceptance and use. For example, the members of the VA Team considered the results of the AIE pilot highly useful. Yet they said they could not explain the methodology to others. If the AIE methodology included ways to foster comprehension and build consensus then it is likely that its acceptance would increase.

It is not clear whether Federal oversight bodies will accept the measurement philosophy of AIE: that the purpose of measurement is to reduce uncertainty using a range or a threshold rather than obtaining a precise number. In time, as more people learn about the challenges and expense of precise measurements, this may be a non-issue, however.

As organizations consider AIE, it may be helpful to think of the automobile, which has become so highly complex that only trained and skilled mechanics using computers can repair it. But this has not kept people from buying and driving cars. On the contrary, there are more cars on the road then ever before. People may come to accept the sophisticated methods of the AIE once they understand that AIE is a tool that provides the means to solve complex investment decisions and to determine the most cost-effective measurement methods.

VIII. Insights into Developing Performance Measures

Observations and discussions during the pilots led to the following insights:

The starting point for development is not the measures but rather a clear understanding of the purpose of the IT investment and its contribution to the organization's objectives.

An IT investment represents a strategy to improve an organization's performance. The closer aligned the investment is to specific organization objectives, the more likely desired performance will improve. Only after the purpose of the IT investment is clear and linked to objectives should performance measures be developed. The pilots clarified the results before discussing the performance measures. More than first half of each pilot was dedicated to defining the contribution of each investment to organizational objectives and clarifying the investment's purpose.

Developing and using performance measures takes time and commitment.

For major IT investments, useful performance measures are the result of in-depth analysis. This takes time and resources. The time invested in developing meaningful performance measures will reduce the time, expense, and loss of momentum from repeated attempts of trial and error during implementation. As measurement data is collected, they must be verified, studied and understood. This requires patience as well as rigor. Once understood, decisions regarding resource allocations and strategy need to be made using measurement information. Changing the way decisions are made not only requires significant changes in management practices but also in mindsets.

Organizational commitment is vital.

An organization's commitment to develop and use performance measures has a direct bearing on the quality of the measures and their implementation. During development, commitment not only involves making key people available, but also involves managers and staff taking time to study and understand the methodology. Commitment during implementation involves making resources available to collect and administer the measures and also finding ways to integrate the results into existing management processes. Without management's commitment, workers will not take measurement seriously. When this happens, workers select measures that are

easy to measure and not necessarily the ones that will provide useful information, which typically are more expensive and difficult to measure.

The implementation of performance measures is just as important as their development.

As challenging as it may be to develop useful performance measures, the proper use of the measures depends on how will they are implemented, who will collect the data, and how the collected information will be used. Devoting time to analyze and decide these matters during development will save time and minimize confusion during implementation. Management must ensure the implementation plans are carried out otherwise they will not be.

Consensus among program personnel and stakeholders is key for acceptance and use.

For major IT investments, taking measurements and making decisions using measurement information will involve many people and substantive changes to an organization's culture. Developing performance measures with input and concurrence from key stakeholders and program personnel greatly increases the likelihood that the performance measures will be adopted and implemented for more than a few months. Until Kaplan and Norton, the developers of the Balanced Scorecard, incorporated consensus building into their methodology many of their corporate clients stopped using their BSCs after six months.

Not all performance measures have the same value.

The pilots revealed that metrics have different values and that some have no value. Each methodology prioritized the items or variables to measure. What's more, there is a trade-off between the accuracy and cost of measurements. For example, surveys have low accuracy because they are based on people's opinions but generally cost much less than controlled experiments that have high accuracy. (Controlled experiments observe the differences between a group that has implemented the IT investment and one or more groups that have not.)

IX. Conclusion

Although the Federal CIO Council completed these pilots in 2001, the performance measures developed and the lessons learned are still relevant for organizations seeking to implement or improve the performance measurement of IT initiatives. The pilots demonstrated that the contributions of a business line (food acquisition and management system) and an infrastructure (security) IT investment to mission and objectives were measurable.

These methodologies link major IT investments to strategic and performance goals. They are useful to federal agencies in the development of their annual performance plans where they must indicate the means and strategies by which the performance goals will be achieved. They are useful to develop performance for the budget exhibit 300s required for major IT investments.

The metrics developed by both methodologies would be compatible with the Performance Reference Model (PRM)[18] released by the Office of Management Budget (OMB) in September 2003 for agencies to use to report performance information for new major IT initiatives. The OMB based the PRM on the Balanced Scorecard and other commercial methodologies. Because both methodologies, particularly the BSC, developed more measures than required for the PRM, a subset of the performance measures would need to be selected to provide the best "line of sight" from the information technology initiative to program and mission outcomes.

The BSC and AIE methodologies differed in their methods as well as their purpose. The BSC methodology built consensus for the top three objectives of five perspectives that represented the strategy to achieve mission results and satisfy customers and users. The purposes for the BSC and its performance measures were to communicate strategies and priorities and to manage the tasks that implement the strategy. The strengths of the BSC methodology were the consensus obtained among senior managers and the scorecard's clarity. A drawback of the BSC methodology is that it did not define methods to collect measurement data. This would impede implementation.

The AIE methodology used scientific methods to analyze the degree that IT security investments would reduce the frequency and severity of IT security related incidents that cause fraud and losses in productivity. The AIE pilot used performance measurements to determine the security investments that would enhance IT security cost effectively. In so doing, a potential $20 million cost avoidance was identified. The strengths of the AIE were the investment analysis that quantified costs, benefits and risks and the measurement methods chosen based upon cost and accuracy. The methods would enhance the efficiency of measurement by reducing their cost. The drawback to the AIE methodology is its sophistication, which requires an analytical background to understand and appreciate. This would impede adoption.

These methodologies are only appropriate for large IT investments because of the upfront cost to contract for methodology experts and the amount of time for government personnel to participate. These methodologies are also best suited for organizations with prior experience with performance measurement and that are committed to improving performance. These methodologies will enhance their management practices and facilitate performance improvement. For other less measurement-experienced organizations or those seeking to develop performance measures only, they should start with logic models[19] that are much easier and less expensive to create and use. Logic models are visual representations of a program's underlying assumptions and causal relationships. They are used for program planning, management, and evaluation.

To benefit from the use of the AIE and BSC methodologies, several ingredients are necessary. An expert with knowledge and experience with each methodology is required. Even the somewhat intuitive balanced scorecard has a subtle sophistication that often goes unnoticed by novices. A person from the host organization is needed to lead the development. They must devote time to learn the methodology, coordinate team members, and promote the findings. This

[18] For more information on the PRM, see http://feapmo.gov/feaprm2.asp.
[19] For more information on logic models, see http://www2.uta.edu/sswmindel/S6324/Class%20Materials/Program%20Evaluation/Executiv.pdf

person should be selected with care because they need to have knowledge of the organization and credibility with senior management. Members of the development team also need knowledge of the organization and the time to develop useful performance measures. Attempting these methodologies without skilled professionals and adequate personnel will not only waste time and resources, their acceptance within the organization will be diminished if not lost. These are not out of the box tools.

The decision as to which methodology to use should be based upon more than just the development of performance measures alone for these methodologies are management tools. The decision also should be based on the need to improve results. To get the most benefit from these tools requires a shift in thinking and changes in management practices that emphasize results not just milestones. This not-so-insignificant change must be accompanied by genuine desire and commitment to improve performance. If the only purpose for doing measurement is to report performance to stakeholders, these methodologies are expensive and time consuming and will become burdensome when used in addition to existing management practices.

Appendix A – BSC Measure Template

The Measure/Target Roadmap (template): Financial

Strategic Objective: F14 Serve More Recipients (Better Use of Dollars for Recipients) **Measure: Increase school participation** **Measurement Intent: Serve more recipients (with same or less dollars by increasing the ability to order and deliver surplus commodities)**	**Frequency of Update: Monthly** **Units of Measure: Percent and/ or dollars**

Measurement Definition/Formula: Current Percentage divided by Total Food Purchase= USDA School Program	
Notes/Assumptions: FNS will provide total food purchase by school district	**Next Steps: Research availability of data**
Measurement Information Is: __X_ Currently Available ___ Available With Minor Changes ___ Not Available	**Data Elements and Sources: FNS, AMS, and FSA (do all agencies have a part of this information? Will the new system be able to track this?)**
Source For and Approach to Setting Targets: Increase school participation percentage (How will you decide what the target will be? Or The baseline will be?	

Target Setting Responsibility: Director, KCCO /Associate Deputy Admin, AMS	**Accountability for Meeting Target: AMS / FSA Divisions**	**Tracking / Reporting Responsibility: AMS/FSA Divisions**	**Measure Availability: July 2002 (school year) Target Avail:**

Target	Baseline Year	2001 Projected	2002	2006
Increase school participation by X %				(Goal for when system is fully implemented)

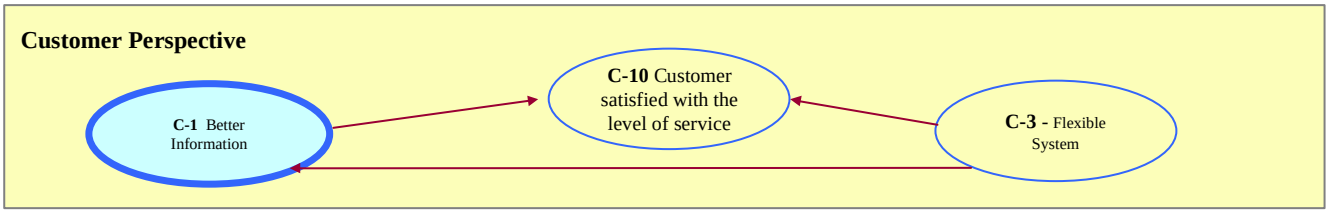
Appendix B – BSC Initiative Template

Strategic Objective: Better Program Delivery (S11)				
Name of initiative: **(S11.A)** Create and coordinate survey to determine non-entitlement school district commodities requirements Initiative accountability: AMS/FNS/FSA Program Divisions				Initiative Start Date: 2003 Expected completion: 2004
Dedicated Resources (financial/asset/people): 5 FSA staff and, 3 AMS				
Anticipated Benefits: USDA supplied products will cost schools less than if procured from other sources thus providing more product for the same dollars or the same product amounts for less dollars.	Total $250,000	2000 projected	2001 projected	2005 projected
Impact on strategic Objectives: C10, C1, S3, L10, F14				
Notes/Issues/Assumptions: Costs include software and establishment of data warehouse.				

Anticipated Progress	Planned Q1	Actual	Planned Q2	Actual	Planned Q3	Actual	Planned Q4	Actual

Appendix C – BSC Pilot Scorecard

FATES – Internal and External Customer Perspective

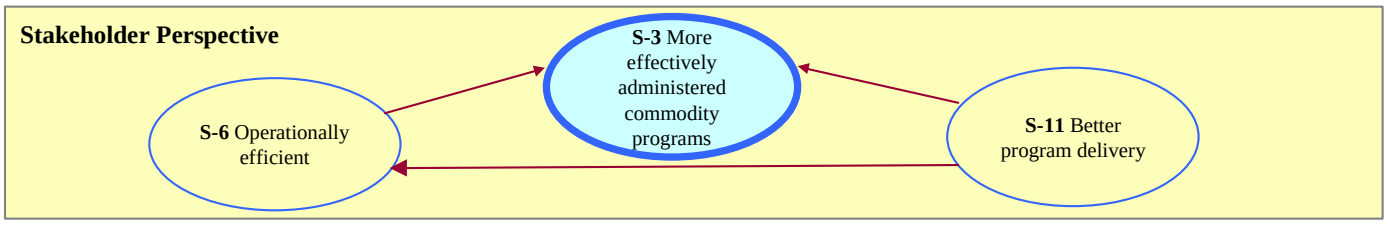


Objectives (Outcomes)	Definitions	Measures	Initiatives
C10 Customer satisfied with the level of service	C10. Provide customer with better sense of control and enhanced feeling of satisfaction through a flexible system that provides better, accurate, timely, information. This will enable us to improve our commodity distribution system.Improve delivery system= reduced time from order to delivery, allow for multiple stops, high seas diversion.	C10.1. Number of calls to Internal Help Desk C10.2.Number of calls to external help desk C10.3 Lead time for users to put in commodity requests. C10.4 Reduce time from survey to delivery C10.5 Internal Customer Survey on Level of Service C10.6 External Customer Survey on Level of Service	C10.A Document time it takes from survey to delivery of commodity C10.B Involve Users in System Development C10.C Create Customer Surveys C10.D Create External User Help Desk C10.E Create Internal User Help Desk
C1 Better Information	C1. Track the data that customers need in real time, eliminate data redundancy, provide easier to retrieve ad hoc reporting; Users will be able to see "their view" rather than searching all data. Data will be reliable and accurate.	C1.1. Decrease in requests to IT staff for reports C1.2. Reliability of data C1.3 Availability of data C1.4 Customer Survey of ad-hoc reporting capabilities C1.5 Accuracy of data	C1.A. Create Survey C1.B. Review COR/PAR/TAR Log for number of current reports requests C1.C. Document incidents of inaccurate data C1.D Document incidents of unavailable data C1.E Correct data accuracy and reliability problems
C-3 Flexible System	C3. Ability to change system to meet program needs in a timely manner without re-writing the whole system (or major parts); Data is reliable and available in timely manner; Ability to acquire and easily use available technology; Responsive to changing business requirements.	C3.1. Decrease of CORS and PARS on CPT Log C3.2. Time to complete system changes	C3.A. Research why system takes so long

Note: Objectives shaded in blue are high impact objectives that affect two or more other objectives within and across perspectives.

Appendix C – BSC Pilot Scorecard (Continued)

FATES - Stakeholder Perspective

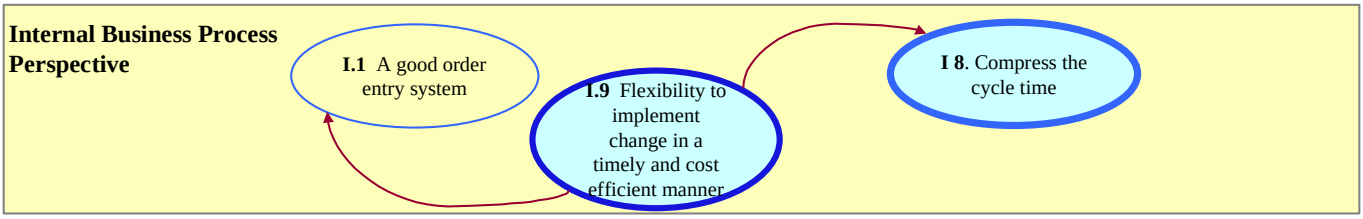


Objectives (Outcomes)	Definitions	Measures	Initiatives
S-3 More effectively administered commodity programs	S3. A universal delivery system; We can use new tools to improve he way we manage our programs. Improve communications (single point of contact) and have communications that drive deeper into customer base (more timely, more targeted) , comply with accounting standards, and provide better funds control	S3.1. Speed, Reduce time responding to queries while reducing the number of total contacts necessary to resolve queries S3.2. Single point of contact S3.3 Number of on-line web site hits S3.5 Operations & Maint. Cost	S3.A Single point of contact (a.k.k. one face for USDA) team S3..B Acquire user friendly, improved, ad hoc reporting system S3.C Develop a survey for users re: manual transactions
S-11 Better program delivery	S11. Deliver more product at same dollar value (on time, when promised) ; Give product promised; Cut down on cycle time; Deliver more programs with same staff.	S11.1.Survey school districts to identify products not available that they are interested in obtaining through USDA S11.2.Number of recipients served per commodity dollar (per dollar allocated and per state state dollar sent to to USDA for purchases through USDA programs S11.3 Number of bid to deliveries within time goals, number of times late	S11.A Create & coordinate survey to determine non-entitlement school district commodities requirements S11.B Policy and Regulatory Change
S-6 Operationally efficient	S6. Cost avoidance (reduction) through the use of better technology (do more with less). Capable of expanding services while holding down operating costs (including system costs and business costs)	S6.1. Enhanced productivity of USDA staff by reducing the resources consumed in performing their functions. S6.2. Cost avoidance through use of modern technology	S6. 1. Provision of new technology tools S6..2. Cost avoidance/ reduction through the use of modern technology

Note: Objectives shaded in blue are high impact objectives that affect two or more other objectives within and across perspectives.

Appendix C – BSC Pilot Scorecard (Continued)

FATES - Internal Business Process Perspective

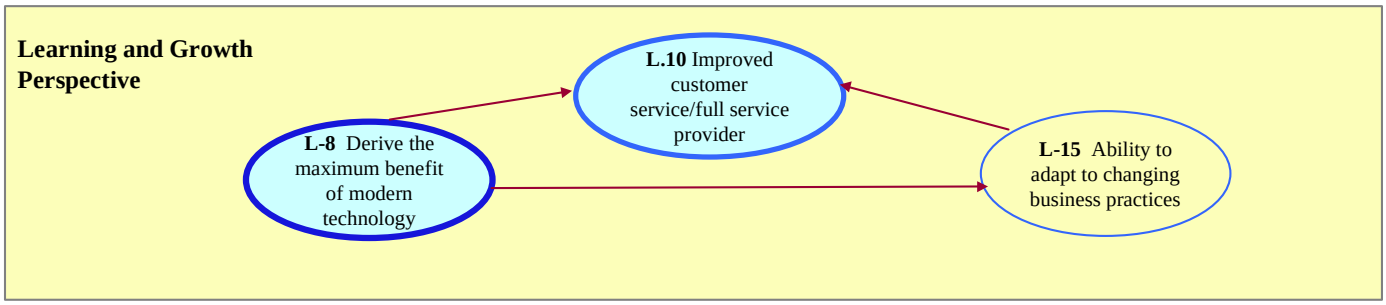


Objectives (Outcomes)	Definitions	Measures	Initiatives
I8. Compress the cycle time	I8. Reduce batch processing (wait time between steps); Data will be better organized to give used "unique" views specified to meet their needs; New functions will allow for feedback Cut down at least X days	I8.1. IT ability to quickly make system changes I8.2 Assessment of time it takes from survey to award of contract I8.3 Compress cycle time from order to delivery by X days	I8.A Procure web-based. Real time COTS
I9. Flexibility to implement change in a timely and cost efficient manner	I9. Allow for the internal system modifications driven by program changes and technology changes; system must be flexible enough to allow for these changes) in timely and inexpensive manner; must allow for system upgrades	I9.1. Current elapsed time to make changes I9.2. Procurement of software development tools I9.3 Procurement of flexible COTS software	I9.A. Procure flexible COTS Software I9.B Procure state of the art development tools
I1. A good order entry system	I1. Accommodate the leap from a few thousand users to possible 30,000; (can't hold 30,000 hands) A good order entry system will help us move bonus commodities quicker . (Real time aspects of the system will allow for additional product demand.;) (determining demand) currently through survey) Allows for faster approval on export side.	I1.1. Requirements and design emphasis on customer perspective I1.2 Customer satisfaction survey I1.3 Reduce time from survey to order	I1.1. Create/ coordinate customer satisfaction

Note: Objectives shaded in blue are high impact objectives that affect two or more other objectives within and across perspectives.

Appendix C – BSC Pilot Scorecard (Continued)

FATES - L&G Perspective

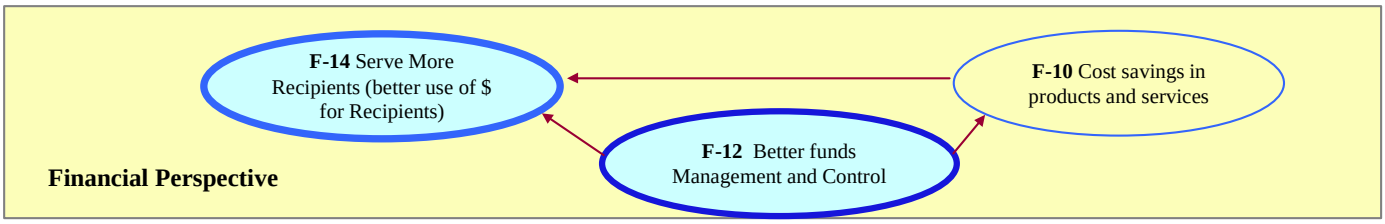


Objectives (Outcomes)	Definitions	Measures	Initiatives
L8. Derive the Maximum Benefit of modern technology	L8. Use web technology to give more users access (train & provide tools) GUI screens (easy to use); Use every aspect of 508; Continuously upgrade technical and business skills in using open systems architecture and new technologies ; Upgrade skills to derive maximum benefit; Employ latest security; Offer "transitional" training; get higher participation in development phase; adapt to commercial practices (leadership promotes changes) ; breakdown barriers between IT and users/ field; (IT-needs better understanding of business process)	L8.1 It self assessment of business practices and new technology skills L.8.2 IT ability to quickly make system changes L8.3 Measure system against security plan during development L8.4 Percent the time for training new users decreases over time L8.5 Measure the system against the security plan after implementation	L8.A. Document current business processes L8..B. Document new technology L8.C Train staff L8.D Internal User self-assessment of training times L8.E Develop Security Plan
L15. Ability to adapt to changing business practices	L15. Better communications (know customer needs; One face on email; automated help desk; People have capability (trained and have tools) to easily react to changing business methods & new technology	L15.1 Users willingness to adapt their business processes to use commercial best practices L15.2 Survey external customers on level of service L15.3 Number of business processes that were changed to CBP	L15A. Implement automated help desk for external customers L.15.B. Train user staff on commercial best practices L15.C Benchmark with companies with commercial best practices L15.D Survey users on knowledge of Commercial best practices
L10. Improved customer service/ full service provider	L10. IT enabled to make changes needed in time that (internal) customers need them; Internal customers enabled to make changes needed in time to allow (external) customers to accomplish their missions.Access to system 24 X 7 (for both internal & external customers); provide same (or competitive) services found in private industry	L10.1. Help desk activity L10.2 Increase in customer satisfaction L10.3 Measure the time it takes to make system changes	L10.A. Survey internal customers on level of service L.10.B.Implement help desk for internal users

Note: Objectives shaded in blue are high impact objectives that affect two or more other objectives within and across perspectives.

Appendix C – BSC Pilot Scorecard (Continued)

FATES - Financial Perspective



Objectives (Outcomes)	Definitions	Measures	Initiatives
F10. Cost savings in products and services	F10. Better return on our program expenditures and system investment; Maximize use of total funding and resources to reduce costs of doing business (e.g. bid lower price due to market changes) (Products= commodities; Services = IT,transportation, warehouse, processing, best value contracting, etc.)	F10.1. Benchmark against private sector (Walmart/ Safeway/Giant) F10.2 Stretch goal of 50% Reduction in maintenance and overhead costs when compared to old system F10.3 Reduced Time from order to delivery	F10.A Benchmark private sector providers and compare their total cost (system and other directly related) to our total cost F10.B. Total process/ system cost reduction F10.C Flowchart complete order entry process from time order is placed through delivery
F14. Serve more recipients (better use of $ for recipients)	F14. Ability to provide more product directly to recipients (more choices of commodities); Able to move more bonus products quicker (right time within right window); Recipients better able to avail themselves of available products; Get products to recipients where,when, and how needed; external customers have direct access to system and information; improved communications; better decision making	F14.1. Number of new commodities offered F14.2 Number of recipients served per commodity dollar spent F14.3 Volume of bonus commodities provided to recipients	F14.A Fully identify the cost of providing products F14.B Customer need identification F14.C Providing Bonus Product Information Faster
F12. Better funds Management and control	F12. Better funds and management control for internal and external customers and agency funds; timely, accurate funds control through FATES (entitlement, obligations, expenditures, allotments and authorizations) so we know how much money has been spent and how much is available by category. ; better decision making.	F12.1. System compliance with all applicable FASAB/JFMIP requirements at acquisition/ as constructed F12.2 System meets all applicable FASAB/JFMIP standards	F12.A. Work with user communities to define and detail requirements F12.B. Work with the financial and program user communities to validate and document funds management and control requirements

Note: Objectives shaded in blue are high impact objectives that affect two or more other objectives within and across perspectives.

Appendix D – AIE Pilot Metrics

Metric: Fraud losses per year

Measurement Method: Continued analysis of reported frauds is critical. Every step should be taken to encourage fraud reporting (emphasize in TEAM). Ultimately, diligent reporting and periodic audits are the best measure of fraud losses.

Metric: Intrusions per year

Measurement Method: Intrusion Detection should report intrusions per year by VA area so that the following can be compared:

- Groups that have been trained under TEAM (Training, Education, Awareness and Message Building) vs. groups that have not
- Groups that rolled out VA-PKI (Public Key Infrastructure) vs. groups that have not
- Groups that have implemented a simplified sign-on solution vs. groups that have not

This is the basis for measuring impact of these initiatives on intrusions per year: "Reduction in Logical Unauthorized Intrusions".

Metric: Pandemic virus events per year

Measurement Method: Anti-virus software should report virus outbreaks by VA area so that groups that have been trained under TEAM vs. groups that have not can be compared. This is the basis for measuring impact of TEAM initiatives on virus outbreaks.

Metrics: Virus productivity impact – specifically:

- Number of VA personnel affected per virus outbreak
- Duration of productivity impact per virus outbreak
- Average annual cost per affected person
- Productivity loss during duration of outbreak

Measurement Method: A random post-event survey of the affected areas should assess each of these (only minor rewording of the current survey is needed). The VIA indicates that a phone survey of 50 to 80 respondents should be sufficient (this should be possible in two days just after the event). Anti-virus reports will also help to put ranges on number affected and duration.