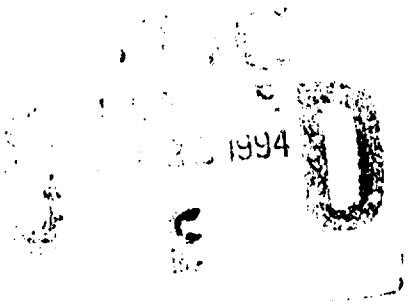
AFIT/GCM/LAS/93S-3

ADA275788




# THE DEVELOPMENT AND DEMONSTRATION OF THE METRIC ASSESSMENT TOOL

THESIS

Cynthia A. Campbell, GS-13, USAF

and

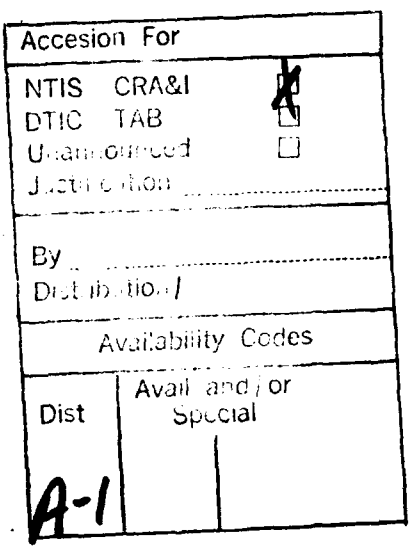
Gregory M. Gutterman, Captain, USAF

AFIT/GCM/LAS/93S-3

165 pg


Approved for public release; distribution unlimited

**The views expressed in this thesis are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.**



AFIT/GCM/LAS/93S-3

# THE DEVELOPMENT AND DEMONSTRATION OF THE METRIC ASSESSMENT TOOL

THESIS

Presented to the Faculty of the School of Logistics and Acquisition Management

of the

Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Contracting Management

and

Master of Science in Systems Management

Cynthia A. Campbell, B.S.
GS-13, USAF

Gregory M. Gutterman, B.S.
Captain, USAF

September 1993

## Preface

The purpose of this thesis was to develop and demonstrate a Metric Assessment Tool for the purpose of evaluating and improving process measurements. We narrowed our study to the inclusion of two demonstrations which supported the viability of the tool. In conducting the demonstrations, we selected individuals who actively participated in the process being measured. We wanted the tool and the improvement actions derived from the demonstration, to be meaningful. We found that the area of metrics is very complex and that developing metrics with proper incentives is a difficult process. Our main hope for this project was develop a tool which will be used to improve metrics and to educate individuals developing metrics on the variety of ways metrics motivate behavior.

We've had a great deal of help during the course of our research. We'd like to extend a specific thanks to Captain Ken Moen at The Armstrong Laboratory for his help and for allowing us to use the human resource laboratory. We'd also like to thank our sponsor, HQ AFMC/XPA, and specifically Captain Mike Pierce, for his aid in identifying our demonstration participants. Finally, we'd like to extend a special thanks to our thesis advisors, Major Kevin Grant and Major Scott Graham, for their expert guidance and remarkable patience over our long journey. They were outstanding.

Cynthia A. Campbell
Gregory M. Gutterman

## Table of Contents

## List of Figures

## List of Tables

**Abstract**

This study reflects the development and demonstration of the Metric Assessment Tool. The purpose of the tool was to provide individuals the means to assess metrics and make improvements to the process measurement. The tool was developed using two critical attributes: customer satisfaction and process improvement. Once the tool was developed, a metric assessment process was designed to demonstrate the tool. Two metrics were selected from the Aeronautical Systems Center (ASC) and individuals with a working knowledge of the metric and process were selected for the demonstrations. Using a Group Support System at Armstrong Laboratory, Wright-Patterson AFB, the group was asked to identify behaviors which might be motivated from the metric. Once the behaviors were identified, the group evaluated the behaviors against the critical attributes. From this assessment, behaviors were placed on the Metric Assessment Tool. This tool clearly identified deficient behaviors and how they might distort the process measurement. From this information, the group was asked to generate improvement actions which would serve to eliminate or control deficient behaviors. With the elimination or control of deficient behaviors, the process measurement is improved and the organizational objective is better served.

# THE DEVELOPMENT AND DEMONSTRATION OF THE METRIC ASSESSMENT TOOL

## I. Introduction

For many years after World War II, the United States dominated the world market in virtually every area. The products produced were in demand and there was little to no competition from other parts of the world. During that era, industries within the United States concentrated their efforts on capacity. Today, the United States can no longer claim preeminence by default. Foreign industries have made great strides into the world market based on their consistent ability to produce high-quality products and services. With the expansion of the industrial base to include foreign nations, the dominance of the United States has declined (4:1-2). To regain its position in the market, the United States has had to meet the challenge of its foreign competitors with a new emphasis on quality.

The Department of Defense (DOD) and, specifically, the Air Force have approached the challenge of quality with a new management philosophy. Total Quality Management (TQM), or Quality Air Force (QAF), was formally adopted in 1988 by then-Secretary of Defense Frank C. Carlucci (8:58). The

essence of TQM lies in continuous process improvement (4:2-8). With continuous process improvement, no practice, policy, tool, system, service, or product is exempt from scrutiny and change. To initiate process improvements and to ensure that process improvements are continuous, clearly defined measures for tracking progress and identifying improvement opportunities are required. Metrics are meaningful measures which present data on the inputs and outputs of a process so that improvement actions can be taken and the effects of these actions verified (5:1). However, the graduate thesis by Miller and Hayes (1992) indicates that many of the metrics developed by the Air Force do not consistently drive appropriate behaviors which result in continuous process improvement (26:5-3). Their findings strongly suggest the need for continuing research focused on the development of effective metrics. A metric helps workers to examine a process to identify process improvement actions. However, there is no tool available to help workers examine the metric itself to identify measurement improvement actions.

**Problem Statement**

A considerable number of metrics have been generated within the Air Force since the inception of TQM. For example, in the Miller-Hayes thesis over 300 metrics were identified from five System Program Offices (SPOs) (26:3-2).

Although significant, this figure reflects only a portion of metrics in existence throughout the Air Force.

The use of metrics is expanding. Where metrics were once generated to ensure the successful accomplishment of organizational objectives, they are now being generated to measure policy compliance. According to a 1992 notification missive issued by General McPeak, Vice Commander of the Air Force, the Air Force will be eliminating its operating instructions and replacing them with policy letters and metrics to measure compliance to the policy (24).

The emphasis on metrics within the Air Force underscores the importance of this research. Metric implementation involves collecting and analyzing data and implementing process change based on that data. As such, metrics consume manpower. Poor quality metrics consume manpower inefficiently. Good metrics are critical to the continuous improvement of Air Force products and services.

Based on the above information, the research team concludes the following problem statement:

> Numerous metrics are being developed within the Air Force yet, there are no tools available to help workers determine if the metrics are effective. Consequently, substantial time and effort are currently being dedicated to develop and use metrics which may not effectively support process improvement, customer satisfaction or the fulfillment of objectives.

The purpose of this research is to develop and demonstrate a means to assess metrics and to identify improvement actions which can be taken to improve the metrics, or the manner in which they are used.

## Research Objective

The objective of this research is to develop and demonstrate a metric assessment tool which will:

1. evaluate metrics based on critical attributes,
2. determine if the behaviors motivated by a metric support the goals and objectives of an organization,
3. identify behaviors which could distort the metric measurement,
4. determine the need for additional supporting metrics, and
5. identify process measurement improvement actions.

## Scope and Limitations

The metric assessment tool will evaluate metrics. Additionally the tool and the evaluation process will identify "process measurement" improvement actions. These actions may go beyond the simple adjustment of the original metric. Recommendations from the evaluation process can reveal the need for a complete metric reconstruction or additional metrics. Therefore, the research team uses the term "process measurement" improvement actions in lieu of "metric" improvement actions to recognize this distinction.

The subject study will have two general limitations. First, the Metric Assessment Tool will be demonstrated on a limited number of metrics. The purview of the thesis study is to "develop and demonstrate" the Metric Assessment Tool.

Therefore, the tool will not be validated under this thesis study. The validation of the Metric Assessment Tool will be recommended for future research in Chapter Five.

Secondly, this research demonstrates the Metric Assessment Tool can be used to identify potential strategies to modify or augment an existing metric, as well as strategies to control undesired behaviors. The study does not investigate whether or not the suggested strategies are actually implemented nor does it confirm that the strategies suggested will successfully improve the process measurement.

## Overview

This chapter introduces one of the most critical elements to the successful implementation of TQM -- metrics. In addition, it discusses the need for a metric assessment tool to evaluate metrics and identify process measurement improvement actions for the Air Force. Also, the chapter outlines the objectives to be accomplished with the assessment tool. Namely, the metric assessment tool will focus on the critical attributes of a proper metric and provide insight into corrective actions which can improve the process measurement. The chapter concludes with a discussion of the scope and limitations of this research.

The remainder of this thesis investigates current research on metrics, describes the methodology used to accomplish the research objectives, discusses the results of

the research applications, and provides conclusions and recommendations for further research. Chapter 2 investigates current research and guidance on metrics and metric development and will discuss the 1992 thesis which identified the need for this research. Chapter 3 will address the method of research, the development of the metric assessment tool, the data collection plan, and the data analysis approach. Chapter 4 will include a review of the metric assessment demonstrations and an analysis of the effectiveness of the metric assessment tool. Chapter 5 will provide conclusions drawn from the research as well as recommendations for any future research that may be required. Appendix A provides a list of definitions which are used throughout this thesis.

# II. Literature Review

## Introduction

In 1981, W. Edward Deming stated that it was management's job to help "people to work smarter, not harder." (10:13) This statement is even more important today. In the United States Air Force, the need to generate quality products and services with limited resources has mandated the adoption of new management philosophies. Total Quality Management (TQM) was formally adopted in 1988 by then-Secretary of Defense, Frank C. Carlucci (8:58). Today, "TQM guides most of what happens" in the United States Air Force (8:58). Total Quality Management is a broad topic with many facets. Therefore, this chapter will focus on one aspect of TQM - quality measurements known as metrics. Quality is driven by process improvements, and process improvements are measured by metrics. Specifically, this chapter will provide an overview of TQM. Next, this chapter will address metrics with an emphasis on the attributes of an effective metric. Finally, this chapter will review prior research concerning metric development and evolution.

## Quality Overview

Total Quality Management is not a new concept. In the late 1800s, Frederick Taylor determined that "a product is

the result of a series of processes that can be studied, improved, and ultimately perfected." (4:3-1) From this idea several concepts were spawned. One of the most notable was a concept introduced by Walter A. Shewart in the 1920s. Shewart "developed techniques to bring industrial processes into what he called `statistical control' (37:6). He emphasized the prevention of defects as an approach to quality. Through the work of dedicated individuals like Deming, statistical process control and other preventive quality tools geared towards systems improvement gained greater exposure. Prior to World War II, Deming taught statistical control philosophies in the United States. However, their use faded during the postwar era when American industries began to value quantity over quality, and when the country's economic supremacy was not challenged by foreign competition (37:12).

Dr. Deming left America and began instructing statistical control processes to Japan in 1950 (10:13). On June 24, 1980, NBC-TV aired a documentary entitled, "If Japan Can...Why Can't We?" (37:11) This documentary sought to discover how Japan had risen from a war torn country after World War II, to an economic giant whose products were far superior in quality to those in the United States (36:11). Dr. Deming was rediscovered.

Today, quality in the United States is receiving increased attention. Major corporations, like the Ford

Motor Company, Xerox, and Chrysler, have all adopted TQM philosophies. Total Quality Management has also found its way into the Federal Government. Former President George Bush stated, "The improvement of quality in products and the improvement of quality in service...are national priorities as never before." (35:Cover) Although America has accepted the idea of quality and TQM, it still follows behind the Japanese. Dr. Deming offers the following analysis of American industries:

> The cause of the decline is that management has walked off the job of management, striving instead for dividends and good performance of the price of the company's stock. A better way to serve the stock-holders would be to stay in business with constant improvement of quality of products and of service, thus to decrease costs, capture markets, provide jobs, and increase dividends. (36:xi)

Total Quality Management is a philosophy that "results in higher-quality, lower-cost products and services that respond faster to the needs of the customer" (31:94). Dr. Deming believes that quality improves productivity and his view is supported by his students. In 1981, Dr. Yoshikasu Tsuda of Rikkyo, University of Tokyo, wrote Dr. Deming and stated:

> I have just spent a year in the northern hemisphere, in twenty-three countries, in which I visited many industrial plants, and talked with many industrialists.
>
> In Europe and in America, people are now more interested in cost of quality and in systems of quality-audit. But in Japan, we are keeping very strong interest to improve quality by using statistical methods which you started in your very first visit to

> Japan....When we improve quality we also improve productivity, just as you told us in 1950 would happen. (10:13)

Dr. Deming believes that management must "constantly improve the system" in order to attain quality in goods and services (10:20).

A colleague of Deming, J. M. Juran, defines quality as freedom from deficiencies and overall product performance. He states, "Deficiencies result in complaints, claims, returns, rework, and other damage." (20:5) Product performance refers to the ability of a product "to be equal or superior to the quality of competing products." (19:4) Thus, the customer defines quality (34:357). Numerous customers, and their differing perceptions of quality, require a system of product development and delivery that is dynamic. Quality, therefore, requires a dynamic system of continual change and improvement, or continuous process improvement. To ensure process improvements have been obtained, measurements must be taken. Measurement, therefore, is critical to the attainment of quality.

Juran's Quality Planning Roadmap includes 10 steps. Step number four recites, "Establish units of measurement to evaluate quality." (20:14)

The Office of Personnel Management for the Federal Government in the United States summarizes its seven operating TQM principles. Measurement is included as principle number four which states that TQM is governed by

"developing clearly defined measures for tracking progress and identifying improvement opportunities." (35:3)

Obtaining quality in the goods and services we provide requires measurement. Measurements establish process control and provide indicators for process improvement.

## Quality Measures (Metrics)

Defining Metrics. A quality measure, process measure, or metric, is defined as follows:

> Metrics are nothing more than meaningful measures. For a measure to be meaningful, however, it must present data that allows us to take appropriate actions. It should be customer oriented and should foster process understanding, thereby motivating action to continually improve the way we do business. (1:5)

A meaningful metric motivates behaviors that are conducive to continuous process improvement. Dennis Kinlaw, a noted measurement consultant, states that "Improvement and measurement always must be kept together as elements in a unified concept." (22:7) H. J. Harrington summarizes this relationship by stating "measurement is the first step that leads to control and eventually to improvement." (15:43)

Metric Importance. An organization cannot realize process improvement, and likewise quality, unless the improvements are being measured (22:7). The Office of Management and Budget (OMB) emphasizes the importance of

metrics in its guide booklet entitled How to Develop Quality Measures That are Useful in Day-to-Day Management.

> One critical element of managing for continuous improvement is to know the level of quality being achieved at any given time and this requires the use of quality measures. Without quality measures, it is entirely possible to be talking about quality improvement while quality is, in fact, declining. (29:3)

Other organizations, such as the United States Air Force, also emphasize the importance of metrics. The ASD Metrics Handbook, a publication of the Aeronautical Systems Division (currently the Aeronautical Systems Center (ASC)), relates:

> Measurement is a fundamental part of good management. Metrics are invaluable from both a program management and a process improvement perspective....Measuring processes provides the basis for appropriate management action(s) to identify opportunities for constructive changes and continuous process improvement. Metrics allow us to baseline where we are, identify the impediments to the process, and track the impact of management actions on processes and other process changes. (1:7)

The Cumberland Group, a quality consulting firm, provides the following description which summarizes the importance of measurement.

> You cannot Manage what you cannot Measure.
> You cannot Measure what you cannot Define.
> You cannot Define what you cannot Understand.
> You cannot Succeed if you cannot Manage! (3:1)

Thus, without a good measurement foundation, organizations cannot expect to attain quality in the goods and services they provide. Metrics are vital to process improvement because they are the only tangible quality indicators. In

1981, Deming reports the following in an article written for National Productivity Review:

> Price has no meaning without a measure of quality being purchased. Without adequate measures of quality, business drifts to the lowest bidder, low quality and high cost being the inevitable result. American industry and the U.S. Government are being rooked by rules that award business to the lowest bidder. (10:18)

Metrics measure and motivate continuous process improvement throughout an organization.

What to Measure. Metrics should motivate continuous process improvement. Managers that set realistic measurement goals will usually achieve those goals. Therefore, the selection of appropriate measures is critical for motivating improvement in the right areas. Deming describes how factories typically perform when improper measures are adopted:

> One will see any day in hundreds of factories, men and women standing around the last hour or two of the day, waiting for the whistle to blow. They have completed their quotas for the day: They may do no more work, and they cannot go home. Is this good for the competitive position of American industry? Those people are unhappy doing nothing. (13:165)

A metric that does not measure quality characteristics will not motivate continuous process improvement. Therefore, the answer to some is to work harder to ensure improvements in the measurement rather than the process. Kinlaw argues that "getting people to work harder in poor systems will produce

minimum improvement." (22:12) Metrics measure (1) output and customer satisfaction, and/or (2) work processes (22:12).

**Measuring Output and Customer Satisfaction.** Measuring customer satisfaction requires the identification of the customer and the customer need. Juran defines the customer as any individual who is affected by a process or product (20:8). He contends there are two kinds of customers: internal and external. Internal customers are those individuals or groups that consume a product or service from within an organization. External customers are those consumers who are outside the organization (20:8). External customers have a critical influence over an organization's reputation. However, internal customers are just as important to quality as external customers. Kinlaw states that "the degree to which work teams satisfy their internal customers will have a profound impact on the organization's ability to satisfy its external customers." (22:98)

The customer defines quality. Customer satisfaction is a "function of quality of service, quality of product, and quality of treatment." (22:101) The United States Air Force provides a service to the country. Thus, quality of service is very important to both internal and external customers. In 1988, the Federal Government conducted a Gallup survey to

determine the characteristics consumers believe provide quality. The survey identified the following characteristics: courteous/polite treatment, promptness, satisfaction of customer needs, and attitude of personnel (28: Attachment A).

Customers, both internal and external, consider themselves satisfied if they have received a service containing these characteristics.

The product received by an internal or external customer is the output of a process. This output has two components: development and delivery (22:31). Customers are interested in output quality. Thus, customer satisfaction is a function of output (22:108). Measuring output allows an organization to answer the following five questions:

1. What is the output costing?
2. How long does it take to produce it?
3. Is the output within our standards?
4. What kinds of errors are occurring?
5. What is the frequency of the errors? (22:32)

Measuring output in order to answer the above questions will allow the organization to begin improving the process of product development and delivery. This, in turn, improves the quality of the output and satisfies the customer.

**Measuring Work Processes.** Customer satisfaction requires process measurements. "Where and how we measure

output depends on the process in which we are interested." (22:32) Kinlaw defines a work process as a "movement or flow of an object (e.g., part, piece of paper, etc.) through a sequence of steps from the point of input to the point of output." (22:129) Large processes may be composed of smaller processes such that the output of one process becomes the input to another.

The purpose of measuring a process is to determine how well a process is functioning and then to find ways to improve the process. The improvement goals for process measures take on three forms:

1. Make the process stable: Ensure that the distribution of the measures taken by the team to determine the performance of the process falls within limits or ranges that should be expected.

2. Reduce the variation in the process: Improve the process so that the distribution of the measures the team takes to determine the performance of the process comes closer and closer together and, therefore, becomes closer to the average of the measure.

3. Improve the average: Move the total process to a higher level of performance so that the average of the measures that the team takes to determine the performance of the process becomes significantly higher or lower (depending on the desired direction. (22:132)

Process measurements are statistically based and typically formatted as a ratio. A full discussion of the statistical aspects of a metric is beyond the scope of this literature review.

Metrics and Teamwork. Metrics provide an organization with a teamwork mentality.

> Measurement creates a higher degree of involvement of people in the goals and processes of improvement. Measurement provides the feedback that helps people to take more active roles in improvement and to take more direct responsibility for it. (22:9)

Continued improvement does not happen independently, it requires organizational teamwork.

> When organizational performance shows significant improvement, this improvement is not the sum of individual performances. It is first of all the result of improved processes. It also is the synergetic result of team performance. (22:12)

Organizations must be aware of the role teamwork plays in continuous improvement. Utilizing metrics to punish or correct employees is detrimental to a team-oriented approach for continuous improvement.

> Continuous improvement will not take hold and survive in work teams or organizations until they are populated by people who are free to ask any question and make any suggestion about any aspect of their work." (22:178)

Thus, teamwork is critical to process improvement.

Section Summary. Metrics are meaningful measures that motivate behaviors that will result in continuous process improvement. The volatile nature of measures requires that organizations measure the right aspects of a process, those aspects which will identify improvement opportunities and drive quality. Lastly, improvement requires teamwork. Without a team effort, process improvement is unlikely. The

next section describes the characteristics of a metric which will properly motivate continuous improvement and likewise quality.

## Attributes of Meaningful Metrics

Section Overview. The importance of metrics cannot be overstated. Measurements will "not always result in process improvement" but, meaningful metrics will (1:5). A meaningful metric results in continuous process improvement and supports organizational goals and objectives (1:5). The following are attributes that the Aeronautical System Center believes a meaningful metric possesses:

1. It is accepted as meaningful to the customer.
2. It tells how well organizational goals and objectives are being met through processes and tasks.
3. It is simple, understandable, logical and repeatable.
4. It shows a trend.
5. It is unambiguously defined.
6. Its data is economically collected.
7. It is timely.
8. It drives the "appropriate action." (5:5)

**Meaningful to the Customer.** The customer defines quality. As discussed earlier, there are internal and external customers. To satisfy the customer, a metric must

take into account what is important to the customer. A metric that does not motivate customer satisfaction, will not drive continuous improvement or quality (22:108).

**Organizational Goals and Objectives Met.** A metric is meaningful if it motivates improvement actions that will ultimately satisfy an organization's goals and/or objectives. (5:1) All metrics must be customer oriented and must be related to or in support of one or more organizational objectives (5:1) Each department within an organization should know how they are contributing separately and together in meeting their strategic mission. Therefore, "measures link operations to strategic goals." (23:7) The Metrics Handbook, a publication of AFMC, states:

> Goals and their subordinate objectives broadly guide the appropriate course of action. In striving to obtain the objectives, you examine your processes, look for ways to improve, and create metrics to track your progress. You should be able to align your metrics to objectives and identify which processes to target for improvement through their application. (5:2)

Currently, AFMC has 5 command goals and 20 command objectives (Appendix B). All metrics within AFMC should relate directly or indirectly to these objectives (5:2).

**Simple, Understandable, Logical and Repeatable.** Metrics must be simple so that individuals can understand their performance and use their experience to improve upon

that performance (22:159). As such, a failure to understand what a metric is saying may prohibit continuous improvement. Juran states:

> An ideal measure is understandable. Many units of measure at the managerial level have involved words that lack standardized meanings or have involved formulas of undue complexity. Any such vagueness or complexity becomes a natural source of divisiveness. Those who lack understanding of the unit of measure become suspicious of those who possess that understanding. (20:76)

Closely related to understanding is the criterion of being logical. According to the Cumberland Group, "if you are trying to accomplish "A", but you are measuring "B", "A" will not happen!" (3:2) A metric must make sense to those who are subject to it because people are much more likely to accept them if they understand why they are important (1:7). There should be a rational connection between the metric and the goals and objectives the organization wants to achieve (1:7). A common problem in this area occurs when too many metrics are being used. The emphasis should be on the quality of the information being collected, not on the quantity of information collected. "Don't try to measure everything. A few good measure is the best policy." (3:2)

In order to show a trend, metrics must be repeatable over time (5:7). The importance of trends will be discussed in the section to follow.

**Shows Trends.** A metric must present data which is useful over time. According to AFMC, "Only trend data has

the potential for evaluating to the necessary degree in order to take action." (5:7) Trend data is used to compare a process before and after improvement initiatives have been implemented (20:77). This allows an organization to gauge its progress towards improved quality. Trend data enables an organization to identify process ailments and improvement actions (5:10).

**Unambiguously Defined.** Juran argues that a good measure "is susceptible to uniform interpretation." (20:77)

> A report on quality of teller transactions in a bank includes number of errors per thousand transactions. Does the failure of a teller to say "thank you" carry the same weight in the report as a key entry error that results in a $500 shortage in a customer account? "Error" must be defined so that its meaning in the report is unambiguous. (20:78)

Because metrics "are widely used as a basis for comparative analysis" (20:76), they must be clearly defined. Recently, organizations have begun to utilize a comparison technique called Benchmarking. Benchmarking is a method used to gauge the performance effectiveness of an organization, or subordinate unit, against the acknowledged quality leader (organization or unit). This cooperation and knowledge-sharing has helped many struggling organizations improve quality (7:19). Both comparative analysis and process understanding require unambiguously defined metrics.

**Economical to Collect.** Juran describes the cost-benefit relationship of metrics.

> It is obvious that a balance must be struck between the cost of making evaluations and the value of having them. In part, the application of this criterion relates to the basic question: Should we measure or not? More usually the application relates to "precision" of measurement. The unit of measure should be established at that level of precision which enables us to make valid decisions from the data. To go beyond that level of precision usually adds cost without adding value. (20:78)

Organizations must weigh the benefits to be gained from metric information against the costs of obtaining the data. The Cumberland Group, a private, quality consulting, firm indicates that organizations should not "try to measure everything." (3:2) A small number of highly meaningful metrics can be effective while keeping the cost of measurement at a minimum.

**Timely.** Metric measures should report a shift or trend in a timely manner so management can act to correct any problems in a reasonable amount of time (12:47). The following example typifies this attribute.

> If a long-term average of only 2% of patients acquire infections in a hospital, and a hospital discharges 500 patients each month, then 10 patients acquire infections per month, on average. If the process worsens (e.g., the true rate goes to 3% infections), it may take several months to detect the shift and effective quality improvement will be impossible. It would be better to measure the component processes of infection control (e.g., intravenous handling and timing) that usually occur every day for every patient. (12:42)

**Appropriate Action.**. This is considered the most important characteristic of a meaningful metric (5:8). Appropriate actions are behaviors displayed by employees that add value to the quality of an organizational product (5:1). The customer defines quality. As such, quality is a dynamic entity. Customers are influenced by a variety of sources, each of which may alter their perception of quality. Therefore, accomplishing quality in the eyes of the consumer demands continuous process improvements (10:20). Actions which are appropriate are behaviors which are motivated by a metric and result in continuous process improvement (23:1). They should be customer-oriented and "should foster process understanding, thereby motivating action to continually improve the way we do business." (1:5)

## Prior Research into Air Force Metrics

The effectiveness of metric measures within the Air Force was the focus of a graduate thesis by USAF Captains Miller and Hayes. Miller and Hayes used a panel of acquisition professionals to analyze a sample of ASC metrics. The acquisition professionals analyzed expected metric response behaviors to determine if these behaviors would result in continuous process improvement. For each metric sampled, a list of behaviors was generated. The likelihood of each behavior and the behavior's contribution to continuous improvement was rated using ordinal scales.

The metric assessment tool used in the Miller and Hayes thesis indicated that many of the metrics evaluated did not foster continuous process improvement. Specifically, Miller and Hayes determined that, "Almost without exception, the behaviors that did not promote continuous improvement" were those associated with measures "that focused on numerical goals or quotas, or otherwise concentrated on meeting the end requirements of the process being measured." (26:5-3) Their findings indicated that, despite the guidance provided within ASC concerning metric development, measures were still being developed that did not promote continuous process improvement. This led Miller and Hayes to recommend future research in metrics (26:5-6).

## Metric Development Processes

Section Overview. Deming stated that "Without guidance, best efforts result in a random walk." (10:12) Because measurement is critical guidance to process improvement, organizations must ensure that tools required to develop meaningful metrics are available. A review of literature reveals that metric development guidance exists within the government. The following subparagraphs describe two approaches to metric development found within the government. The first is guidance from the Office of Management and Budget. The second is from the Air Force

Materiel Command. Both methods lack substantial information concerning the assessment of metric effectiveness.

**Office of Management and Budget.** The Federal Quality Institute, under the Office of Management and Budget (OMB), publishes a metric guide entitled, How To Develop Quality Measures That Are Useful In Day to Day Management (January, 1989). This guide provides three methods for metric development. Method One is the most in-depth. The following is a summary of this method.

1. Identify all customers of the program's outputs and those customers' requirements and expectations.
2. Define the entire work process that provides the product/service.
3. Define the value-adding activities and outputs that comprise the system.
4. Develop quality measures or indicators.
5. Assess quality measures. (29:10-16)

Step five, assessing quality measures, is further described as follows.

> To be sure they will be useful, evaluate the measures that are initially proposed using the following criteria:
>
> a) Are they formulated at critical points in the total work process, i.e., at steps in the process where value-adding activities produce intermediate and final outputs?
>
> b) Do they encompass a controllable activity? Since the intent is to use this information to verify and make improvements, it is important that the measure is able to reflect any action taken to change the process.

c) Is it feasible to obtain, in a regular manner, the data needed for each measure?

d) Have the users of the measures been identified and their needs incorporated?

e) Have descriptive terms (e.g. thorough, consistent, accurate) been clearly defined?" (29:16)

These questions relate directly to some of the attributes of meaningful measures discussed earlier in this chapter.

**The Metrics Handbook.** This guide is utilized by a variety of Air Force Materiel Command units. The handbook provides a 10-step process for metric development. For brevity, the steps that describe metric data presentation are not included.

1. Identify your purpose.
2. Develop your operational definition starting point.
3. Identify and examine existing measurement systems.
4. Generate new metrics if existing metrics are inadequate.
5. Rate your metric against the "eight attributes of a good metric. (5:9)

Step 4 identifies the need for metric assessment. The handbook's explanation of this step is as follows.

> Most measurements used in the past were not process oriented. They were results indicators related to final outputs, products or services for external customers. With metrics, the focus is on how processes are performing in making these final outputs. We are interested in those upstream process measures which derive the final outcome and are the key to making process improvements. The assumption is: if process

> performance is monitored and improved, the quality of the products and services will improve. (5:9)

The document stipulates the need for metric assessment but fails to include guidance on how metrics are to be assessed.

## Summary

Currently, there exists a plethora of information concerning quality measures, but little information on metric assessment. Miller and Hayes found numerous ASC metrics to be deficient in their ability to foster continuous improvement. Without a defined metric assessment technique, the Air Force will continue its ineffective use of metric implementation manpower. Clearly, ineffective metrics stagnate our ability to improve quality and quality improvement is crucial to the economic survival of this country.

# III. Methodology

## Introduction

The objective of this research was to develop and demonstrate a metric assessment tool which would evaluate behaviors generated by current metrics and identify process measurement improvement actions. In order to specifically address the research objective identified in chapter one, the research team developed a six-step methodology. First, the research team conducted a literature review to explore metrics and to determine the critical attributes of an effective metric. The team used this information to establish the metric evaluation criteria and to develop the metric assessment tool. Second, the B-2 System Program Office (SPO) and the Science and Technology (S&T) laboratory office were identified as participants for the metric assessment. Third, one metric from each participating office was selected based on a mutual agreement between the office, the sponsor, and the research team. Fourth, a metric evaluation group was identified from each participating office. The metric evaluation group was composed of individuals who were intimately involved with the process measured by the selected metric. Fifth, the evaluation group assessed the selected metric using the metric assessment tool and a variant of the Nominal Group Technique. A Group Support System (GSS) provided by

Armstrong Laboratory (AFMC/AL/HRG) facilitated the assessment demonstration. Finally, the evaluation group was asked to comment on the value of the metric assessment tool and the metric assessment process in a questionnaire developed by the research team.

## The Metric Assessment Tool Development

Critical Attributes. As described in Chapter II, the Air Force Materiel Command (AFMC) identified eight attributes of a good metric. Based on an analysis of these attributes, the research team classified the attributes into two categories: those which addressed the "structure" of a metric and those which addressed the "purpose" of a metric. The metric assessment tool was designed to assess metrics and identify process measurement improvement actions. The improvement actions may, or may not, alter the original metric structure. Therefore, the attributes addressing the "purpose" of a metric were considered "critical attributes" for the subject research. Table 3.1 summarizes the critical attributes.

The Metric Assessment Tool. The metric assessment tool evaluates metrics and identifies areas where actions can be taken to improve the process measurement. The tool was developed using two of the three critical attributes identified in Table 3.1: meaningful to the customer

**TABLE 3.1**

**CRITICAL ATTRIBUTE SUMMARY**

| CRITICAL ATTRIBUTES | EXPLANATION |
|---|---|
| **MEANINGFUL TO THE CUSTOMER** | A metric must generate behaviors which will meet or exceed a customer's requirements. As such, it must support **"customer satisfaction"**. |
| **OBJECTIVE SATISFACTION** | A metric must generate behaviors which support progress toward the fulfillment of an organization's goals and objectives. |
| **APPROPRIATE ACTIONS** | A metric must motivate behaviors which will foster continuous process improvements. In other words, the metric should motivate **"improvement actions"**. |

(customer satisfaction) and appropriate actions (improvement actions). These attributes were used as criteria for assigning behaviors generated by the evaluation group to the metric assessment tool. The positive and negative aspects of the attributes define the four quadrants represented in Figure 3.1.

The assignment of behaviors to quadrants identifies measurement deficiencies and provides insight into potential process measurement improvement actions. An explanation of the measurement deficiencies associated with each quadrant follows Figure 3.1.

The remaining critical attribute, objective satisfaction, was assessed in two ways; first, by examining the composite of behaviors in all four quadrants and second, by examining the behaviors in quadrant A only. Both the composite behaviors and quadrant A behaviors were evaluated using ordinal scales. Together, these evaluations indicated the extent to which metrics support an organizational objective both before and after process measurement improvements. Further details regarding the objective evaluation are provided in the Group Evaluation Process section of this chapter.

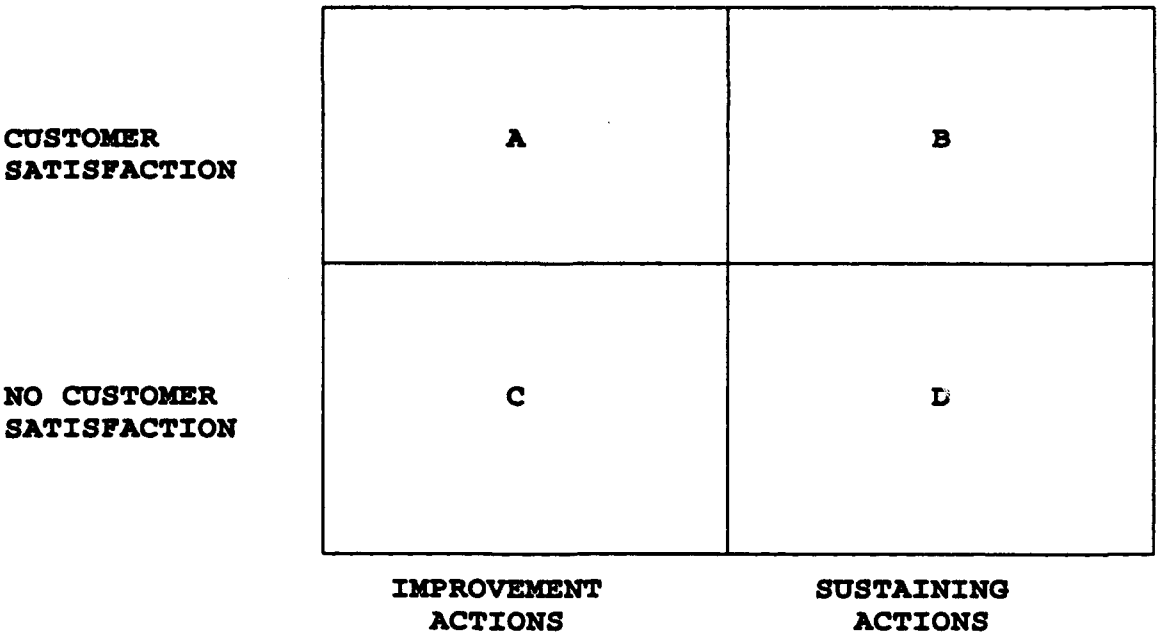



**Figure 3.1 The Metric Assessment Tool**

**Quadrant A.** Behaviors that fall within Quadrant A reflect a motivation towards process improvements which will

result in customer satisfaction. As such, these behaviors reflect the ultimate goal of a good metric (i.e., improving the process and customer satisfaction). Quadrant A also provides insight into whether or not the behaviors generated by the metric support an organization's objective. For example, if the organization's objective was to decrease the number of defective parts produced and a behavior identified in Quadrant A was to emphasize inspection at critical inspection points, one might contend that the metric, and thus the process measurement, are promoting the objective.

**Quadrant B.** Behaviors that fall within Quadrant B are behaviors which are not directed towards process improvements. They are generally temporary actions done as an immediate response to ensure customer satisfaction or to meet some particular measurement goal. These behaviors may ultimately distort the process measurement. For example, if a measured process indicates that 50 contracts were awarded within 100 days. The customer would be satisfied because he has received his award. However, a closer look into the process reveals that people stayed until midnight in order to award 5 of the 50 contracts. The process measurement is distorted because it no longer reflects the actual results of the process. The process only generated 45 contracts, but the extra hours made it appear as though the process was generating more.

**Quadrant C.** Behaviors that fall within Quadrant C are behaviors which are directed towards process improvement but do not support customer satisfaction. For example, if a metric motivates a process change for employees to account for their efforts on a per-project level, the process change may not influence customer satisfaction. As a result, the behaviors would not necessarily support the delivery of quality products and services and may not be considered a valued improvement.

**Quadrant D.** Behaviors falling in Quadrant D are similar to the behaviors in Quadrant B in that they may ultimately distort the process measurement. In addition, behaviors falling in Quadrant D will not support customer satisfaction. For example, a customer wants delivery in 30 days. The metric measures the actual delivery days to the scheduled delivery days as a ratio. One behavior which might be generated by this metric would be a modification of the scheduled delivery date to the customer. Although the customer may not have a choice but to agree to this new date, he will probably be dissatisfied. The behavior did not result in customer satisfaction, nor did it motivate process improvement.

**Activity Selection**

ASC has been acknowledged as the Air Force Materiel Command (AFMC) leader in the implementation of Total Quality Management (2:iii). Based on this fact, and the economical convenience of its location, the research team selected organizations within ASC for the research effort. It should be noted however, that the metric assessment tool is believed to be applicable to any organization involved in the development and/or application of metrics. The research team established three criteria for the identification of the ASC offices who would participate in the metric assessment. First, an office must have metrics which apply to processes performed within that office. This was to ensure that the evaluation group selected from the office had a working knowledge of the process and behaviors which might occur within the process. Second, the office must be willing and able to commit its resources for the metric assessment. The research team required five individuals from each office for a four- to six-hour period. Finally, the research team required that the selected office be able to link its metrics with HQ AFMC command objectives.

Metrics are the diagnostics which measure progress in meeting a command's goals and objectives. During the Bush Administration, the following vision of the Air Force was adopted:

> Air Force people building the world's most respected Air and Space Force...global reach and global power (16:4)

Supporting this vision, the Air Force Materiel Command (AFMC) developed the following mission statement:

> Through integrated management of research, development, test, acquisition, and support, we advance and use technology to acquire and sustain superior systems in partnership with our customer. We perform continuous product and process improvement throughout the life cycle. As an integral part of the Air Force War Fighting Team, we contribute to affordable combat superiority, readiness and sustainability. (16:4)

Advocating the mission statement, 5 command goals and 20 command objectives were developed. The AFMC command goals and objectives are outlined in Appendix B. For each of the AFMC objectives, specific metrics were developed to ensure that the objectives were accomplished. However, all metrics must directly or indirectly support an organization's goals and objectives (23:5). The objectives ensure that all components with the Air Force are moving in the same general direction. AFMC metrics and processes which do not support it's goals and objectives are questionable (17:3).

Based on the above criteria, the B-2 System Program Office (SPO) and the Science and Technology (S&T) laboratory office were selected by mutual agreement between the offices and the research team. The selection of these offices allowed the research team to demonstrate the applicability of the Metric Assessment Tool.

**Metrics Selection**

The metric assessment tool was designed to apply to any type of metric. Two criteria were established for the selection of metrics. First, the metrics selected must apply to a process which was performed within the participating office. This was to ensure that the individuals selected for the evaluation group had a working knowledge of the metric and the behaviors which might, or have, occurred as a result of the metric. Second, the metrics should be under the authority of the participating office. In other words, the participating office should have full responsibility for metric enhancements and process control. The measurement improvement actions generated from this research are more meaningful to those who are responsible for their implementation. This metric selection criteria was developed to enhance the quality of the suggested improvement actions generated by the evaluation team.

Based on the above criteria, and in coordination with the sponsor, the B-2 and S&T offices selected one metric each to be assessed. Two metrics, in total, were considered sufficient for the metric assessment demonstration. Once the metrics were selected, the B-2 and S&T focal points identified the command objectives supported by the metric.

**Group Selection**

The research team solicited the assistance of the B-2 and S&T focal points to identify the metric evaluation group. Two selection criteria were established. First, the individuals must be involved in the day-to-day operation of the process. Noted author Carl M. Moore asserts, "if you want to affect policy, it is wise to include those responsible for acting on the policy." (27:16) These individuals were selected because they had a working knowledge of the process and the behaviors which might occur within the process as a result of a metric.

Secondly, individuals selected for the evaluation group must have completed some formal training in TQM. Selected individuals needed a practical understanding of what constitutes a process, the purpose of metrics, and the meaning of continuous process improvement.

To determine the number of individuals required for the metric assessment, two constraints were considered. First, the metric evaluation process was to be implemented using the Group Support System (GSS) located at the Armstrong Laboratory, Wright Patterson AFB, Ohio. The number of computer stations available at this facility constrains the maximum group size to eight. Secondly, the number of resources taken from the SPOs at one time had to be minimized.

The optimal size of a decision-making group varies. Although larger groups have greater pooled intelligence for problem-solving, they also have more difficulty communicating and reaching a consensus (14:214). Conversely, if a group is too small, its ability to effectively generate many alternatives diminishes. According to several studies, a group size of five is very effective for small discussion groups involved in decision-making (14:214, .19:86, 25:835). The selection of five individuals from the participating offices would not overly burden the SPOs selected for the research and would still produce high quality decisions. Based on the number of individuals required and the established selection criteria, the B-2 and S&T focal points selected the individuals for the evaluation.

## The Metric Assessment

The Assessment Technique. The selected evaluation groups utilized the following technique to assess metrics.

**Group Decision-Making.** A group decision-making process was used to evaluate the metrics selected for evaluation. Studies indicate that groups yield more accurate decisions than individuals alone (14:329, 19:99).

In fact, according to one study, groups outperformed their most proficient group member 97 percent of the time (25:834).

The Nominal Group Technique will be used to implement group decision-making.

> The Nominal Group Technique (NGT) is a method for structuring small group meetings that allows individual judgments to be effectively pooled and used in situations in which uncertainty or disagreement exists about the nature of a problem and possible solutions. (27:24)

According to recent studies, the NGT is considered to be superior to other group decision-making techniques in terms of decision accuracy and/or quality (19:98). The NGT typically includes four steps:

1. Silent and independent generation of ideas by individuals in the presence of a group.
2. The presentation of ideas to the group without discussions.
3. A serial discussion of ideas for clarification and elaboration.
4. The silent and independent ranking of ideas for the final decision. (19:98, 27:24)

For the subject research, potential behaviors which may result from a metric were generated independently by the group and simultaneously presented to the group using the Group Support System (GSS) software. After the behaviors were generated, open discussions were held and questions encouraged. Dialogue was limited to clarifications and the consolidation of duplicate ideas. Where necessary, the wording of an idea was modified upon the mutual agreement of the group. Finally, the group rated each behavior in

accordance with the metric assessment tool parameters (i.e., customer satisfaction and process improvements) in order to assign the behaviors to the tool quadrants. A more elaborate discussion of the evaluation process is provided in the Group Evaluation Process section of this chapter.

**Group Support System.** The Nominal Group Technique was implemented using a Group Support System (GSS). A GSS is an interactive computer-based system which combines communication, computer, and decision-making methodologies to support the formulation and solution of unstructured problems by a group (18:266). The GSS allowed group members to work under anonymous and dispersed conditions. According to a University of Arizona study (1988), members working under anonymous and dispersed conditions generated more comments and engaged in more thorough treatment of alternatives than members using nonautomated group problem-solving techniques (18:276).

> In part, these systems (GSS) are electronic implementations of older methods - e.g. Delphi and Nominal Group Technique - that have been used to improve the quality of meetings over the last 30 years....Over 88% of the users in these studies felt the system had improved the quality of the decisions reached. (6)

The GSS used for this research was located in the Armstrong Laboratory, Wright-Patterson AFB, Ohio. The Armstrong Laboratory primarily performs research and development which

focuses on technologies for improving the performance of people, information, and equipment in the functions of acquisition and logistics support (9:1).

The software used in the GSS was the GroupSystems V. The GroupSystems V is commercial software developed by the University of Arizona (funded by IBM TeamFocus) (30:6). The software implemented the NGT and provided a report of all decisions and comments generated by the evaluation group. In the context of this study, the software recorded and displayed the behaviors generated by the group, allowed the group to evaluate the behaviors against the established criteria, and provided instant feedback on the results. This instant feedback was used to generate the process measurement improvement actions. One of the notable features of this software was its ability to adapt to the specific needs of the research team and the evaluation group. Screens were customized to support each phase of the metric assessment process.

Assessment Process Verification. The Armstrong Laboratory personnel worked closely with the research team to adapt the metric assessment tool and the evaluation criteria to the GSS. Prior to the actual metric evaluations, the research team participated in a software orientation session and two pretests in order to determine

the group assessment process to be used in the demonstrations.

**Software Orientation.** The objective of the software orientation was to familiarize the research team with the GroupSystems V software used to implement the group evaluation process. Personnel from Armstrong Laboratory demonstrated the ability of the GroupSystems V to capture ideas while simultaneously displaying the ideas of others. In addition, laboratory personnel demonstrated the software's editing capability and how the software would automatically assign behaviors to the metric assessment tool based on the group's response to the evaluation criteria.

**Pretest I.** Once the software programs were customized and the orientation complete, the research team conducted the first pretest of the proposed group evaluation process. The objective of this pretest was to define the process in full detail. The research team went through each phase of the metric assessment process to ensure that the process was complete and ready for the second pretest. Each screen and all instructions were evaluated to ensure accuracy and clarity.

**Pretest II.** Any changes resulting from the initial pretest were incorporated prior to the second

pretest. In the second pretest, a mock evaluation group evaluated the following metric selected from the Miller and Hayes thesis (26:4-1).

> The Number of Undefinitized Contractual Actions (UCAs) definitized within 180 days vs. those definitized after 180 days.

The metrics used by Miller and Hayes were derived so that they would be generic to many organizations within ASC. As such, a specific command objective was not available. Therefore, the research team reviewed the command objectives and selected the following objective for Pretest II:

> Objective 1.3: Be our customers' supplier by choice by: meeting cost, schedule and performance baselines, enhancing customer support and lowering life cycle costs.

The research team believed that the metric was designed to encourage the timely definitization of contract actions to benefit the customer. During Pretest II, the evaluation group assessed the metric in accordance with the pretest metric assessment process identified in Table 3.2.

To aid in the assessment, three handouts were provided to the pretest evaluation group. First, the group was given an outline of the process as well as examples of the types of responses needed for each phase (Appendix C). Second, the group was given a description of the metric assessment tool and an interpretation of each quadrant to aid in the research team's explanation of the tool (Appendix K). Third, the team was provided with a software help guide

(Appendix D as revised) which briefly outlined critical software commands or keys to be used for implementing the assessment process. During the pretest, content validity was ratified by using the critical attributes of a good metric and by receiving confirmation of the reality

**TABLE 3.2**

**SUMMARY OF PRETEST METRIC ASSESSMENT PROCESS**

| STEP: | ACTION: |
|---|---|
| ONE | IDENTIFICATION OF BEHAVIORS |
| TWO | DISCUSSIONS/REFINE BEHAVIORS |
| THREE | IDENTIFY CUSTOMER |
| FOUR | CRITERIA EVALUATION |
| FIVE | EXPLANATION OF METRIC ASSESSMENT TOOL |
| SIX | EVALUATION OF ORGANIZATIONAL OBJECTIVE |
| SEVEN | BRAINSTORM IMPROVEMENT ACTIONS |
| EIGHT | DISCUSS/REFINE IMPROVEMENTS |

behaviors in the questionnaire (Appendix L). Reliability was enhanced by the establishment of consistent instructions which were read verbatim to each evaluation group. As a result of this pretest, seven action items were identified by the pretest evaluation group.

*Pretest II Action Items.* Seven action items were generated as a result of the second pretest. The

action items, as well as a complete explanation of each item, are as follows:

1. The Introduction. Provide a more thorough explanation of the metric assessment process with examples.

2. The Introduction. Eliminate the reference to outcomes versus behaviors.

3. Step Three: Identify Customer. Clearly define "primary customer." Also, redefine Step Three as "Voting on Primary Customer." Obtain a list of potential customers prior to the metric assessment and allow the evaluation group to rank order the list.

4. Step Four: Criteria Evaluation. Change the evaluating structure from a Yes/No scale to an ordinal scale.

5. Step Seven: Improvement Actions. Allow the evaluation group to identify improvement actions for individual behaviors, metric structure and others as three separate actions. Eliminate the discussions and allow the research team to generalize the results.

6. Step Eight: Eliminate discussions on improvement actions.

7. Handout (Appendix D). Provide a more detailed software help guide so that participants can more readily understand the data input procedures for each step in the assessment process.

During Pretest II, some of the participants identified the need for a more thorough presentation of the metric assessment process. They indicated that the examples provided in the first handout would have been more beneficial if they had been briefed along with the process explanation. Also, there seemed to be some confusion over the distinction made between outcomes and behaviors in Step

One (identification of behaviors). Therefore, in order to avoid confusion, the research team eliminated all references to outcomes from the instructions to Step One. In Step Three, the research team asked the evaluation group to identify internal and external customers who might benefit from the output of the process being measured. After much discussion, the research team found that there was some confusion over whether an entity was a customer of the output of the process or a customer of the product which the process supports. For example, the process evaluated in the pretest provides a signed UCA as an output. This process provides a new contractual document to the customers in the SPO who are trying to procure a new military product. Alternatively, the customer might also be the users in the MAJCOM who will actually receive the new product. The evaluation group decided to select the users in the MAJCOM as the primary customer. However, it was obvious that the research team needed to define the primary customer more clearly. In addition, the research team found that the time required to identify and clarify potential customers could be eliminated by obtaining a customer list prior to the metric assessment. To accomplish this, the research team solicited potential customers from the participating organization's focal points.

In Step Four, the research team asked the evaluation group to answer two questions for each behavior regarding

customer satisfaction and process improvement. According to the software selected for this phase, the groups answers were to be either a "yes" or a "no." During the pretest the evaluation group found that many behaviors could not be defined in terms of a definite yes or no. Therefore, the research team decided to modify the questions so they could be answered with an ordinal response scale.

In Step Seven, the research team asked the evaluation group to identify improvement actions which might be done to control behaviors, improve the metric structure, and support the metric all at once. The group indicated that it was too difficult to change their thought process from structure actions to behavior actions and back again. More structure was required. Therefore, in the final assessment process the evaluation group addressed improvement actions for each individual behavior under Phase I, the metric structure under Phase II, and any other actions under Phase III.

In Step Eight, the research team had contemplated the use of a discussion period to allow the groups to fully explain their improvement actions. However, during the pretest it was agreed that the most critical purpose of the process was to generate as many improvement actions as possible in the time available. Discussions of each individual recommendation would have been exhausting for the evaluation group and might have eliminated viable options for the manager receiving this information. Therefore, it

was decided that the group would generate the improvement actions and the research team would generalize the results into categories for the manager to use according to his/her own needs.

The final pretest action item concerns the software handout. A total of four software applications were used at various times in the assessment process. Table 3.3 identifies the software application to its applicable step in the process. Each software application has its own procedures for inputting data or casting votes. For example, in the Categorizer, an insert key must be hit before data can be entered. In the Alternative Evaluator application, an enter key must be hit before data is

**TABLE 3.3**

**SOFTWARE APPLICATIONS**

| PROCESS STEP | ACTION | SOFTWARE APPLICATION |
|---|---|---|
| ONE | Identify Behaviors | Categorizer |
| THREE | Primary Customer Voting | Rank Order |
| FOUR | Criteria Evaluation | Alternative Evaluator |
| SIX | Organizational Objective Evaluation | Ten Point Scale |
| SEVEN | Identify Improvement Actions | Categorizer |

entered. In the Alternative Evaluator application, an enter key must be hit before data is entered.. The software instructions provided to the group lacked this type of specific guidance. Therefore, it was apparent that clearly defined software instructions for each step in the process were required.

Improvements identified as a result of Pretest II were implemented and established the metric assessment process to be used in the demonstrations.

The Metric Assessment Process. In the assessment process, each evaluation group met from four to six hours to evaluate one metric. The evaluation consisted of an eight-step process preceded by a process orientation session. The assessment process is summarized in Table 3.4.
Two metric assessments were demonstrated (S&T and B-2) in this research project. The process to follow includes adjustments made based on lessons learned from the first demonstration (S&T). Final process adjustments will be discussed in Chapter 5 based upon the research team's final conclusions.

**Process Orientation.** The purpose of the process orientation session was to familiarize group members with the software and equipment they would be using for the evaluation. To facilitate this session, the metric

**TABLE 3.4**

**SUMMARY OF METRIC ASSESSMENT PROCESS**

| STEP: | ACTION: |
|---|---|
| ONE | IDENTIFICATION OF BEHAVIORS |
| TWO | DISCUSSIONS/REFINE BEHAVIORS |
| THREE | PRIMARY CUSTOMER VOTING |
| FOUR | CRITERIA EVALUATION |
| FIVE | EXPLANATION OF METRIC ASSESSMENT TOOL |
| SIX | EVALUATION OF ORGANIZATIONAL OBJECTIVE |
| SEVEN | IDENTIFICATION OF IMPROVEMENT ACTIONS |
| EIGHT | DISCUSS/REFINE IMPROVEMENTS |

evaluation group was provided with a software reference guide (see Appendix D). In addition, a sample metric was used to demonstrate each step of the evaluation process (see Appendix E). One member of the research team acted as the group facilitator. The facilitator provided the instructions for each step in the process and answered any questions from the group.

The process orientation was used to eliminate any anxieties the evaluation group may have had regarding the days activities. During this session, the participants were informed that they could withdraw from the demonstration at any time throughout the evaluation process if they felt uncomfortable about continuing. Prior to the start of step one in the process, the group was provided with a Metric

Assessment Fact Sheet, Appendix G (Appendix F for S&T group), which outlined the metric to be evaluated as well as the organizational objective. (Final time: 10 to 15 minutes)

**Step One: Identification of Behaviors.** In Step One of the assessment process, the evaluation group was asked to independently identify behaviors which may result from the metric under evaluation. The following question was proposed to the evaluation group:

> What behaviors might occur as a result of this metric?

The evaluation group was then given the following instructions:

1. No discussions or value judgements were to be made during this step of the process. This was to be a silent brainstorming session.

2. A discussion of the "it depends" syndrome. The group was asked to include both positive and negative behaviors and to make clear the intent or impact of the behavior to the process. In other words, a behavior could have both a positive and negative connotation "depending" on the intent of the author. The group was told that they should address their ideas as completely as possible to include the negative and positive implications of their thoughts. They were also told not to concentrate on this aspect too much because any behaviors which might be unclear would be addressed in discussions under Step Two. The research team wanted them to be aware of the "it depends" syndrome but did not want it to hinder their imaginations. The most critical issue was to identify and register ideas.

3. If the behavior identified was a specific action, the actor must also be identified.

4. The ideas generated should be behaviors not outcomes. In the first demonstration (S&T Demonstration), the facilitator did not reference the difference between outcomes and behaviors because the pretest indicated that it was confusing. However, the behaviors generated in the first demonstration indicated a need to readdress this issue. Therefore, in the second demonstration (B-2 Demonstration), the facilitator made it clear that the assessment process required the generation of behaviors not outcomes. To ensure this distinction was clearly understood, the facilitator provided Appendix H to each participant and used the examples therein.

Each group member had access to an individually dedicated computer terminal. Although the group members worked independently, their ideas were anonymously recorded onto a master list which was displayed on their terminals and on a large wall screen in the front of the room. The purpose of the master list was to foster creativity by eliciting new ideas from the ideas of others. In addition to the prior instructions, the facilitator also provided software instructions. The following instructions were provided:

1. The behaviors to be identified were to be limited to one line. However, additional comments or explanations could be included in the comment section of their screens.

2. The group should not be concerned with punctuation, spelling, or grammar while generating their ideas. Corrections would be made in Step Two.

3. The group was told that there was no time constraint so they should generate as many ideas as possible.

4. The group should not be concerned about duplications. Duplications would be consolidated under Step Two.

5. The facilitator referenced Attachment D, The Software Help Guide, to begin Step One.

The outcome of Step One was an unedited list of possible behaviors that might be driven as a result of the metric under evaluation. (Software Application: Categorizer. Final time: 20 to 30 minutes) Note: All time estimations are contingent upon the number of behaviors generated. The times used in this chapter are based on an average of 18 behaviors.

**Step Two: Behavior Discussions and Refinements.** The purpose of Step Two was to clarify and refine the list of behaviors developed in step one through group discussions. The facilitator proceeded down the list, one behavior at a time, asking group members to comment on any behaviors that were unclear to them. The software allowed access to any additional comments provided by the author for further explanations. Any duplications found were placed into a software file behind the original idea. The intent of the discussions was to foster a better understanding of the behaviors generated. Any behaviors which were not understood by the group were either clarified or eliminated by mutual agreement. To ensure that each group member had a say in the final outcome, each member was individually asked if they agreed with the final version of the behavior. Any behaviors that were viewed more as outcomes were revised or eliminated. In addition, as each behavior was reviewed the facilitator asked the group members to ensure that the

impact or the intent of the behavior was clear. To ensure that the group had a clear understanding of the "it depends" syndrome, the group was referenced to Appendix H for an example (also discussed in Step One). If the impact or intent of a behavior was unclear, the behavior was either reworded to express the implication or was split into two ideas, the first carrying the positive connotation and the second carrying the negative connotation. The product of Step Two was a list of collectively exhaustive and mutually exclusive behaviors which were understood by all members of the group. (Final time: 30 to 40 minutes)

**Step Three: Voting on Primary Customer.** In order to proceed with Step Four a *primary customer* must be identified so that all members of the group would be identifying with the same entity for the customer satisfaction criteria evaluated in Step Four. Prior to the beginning of the session, the office focal point provided the research team with a list of potential customers. The research team included an "other" category along with this list to enable group members to add other customers as needed. From this list, the group was asked to rank order the customers from the most primary customer to the least primary customer. A primary customer was defined as follows:

> The customer who benefits the most, and most directly, from the output of the process the metric measures.

The research team wanted to capture the customer who would benefit the most from process improvements. The customer receiving the greatest number of first place votes was considered the primary customer for the metric assessment. If no single customer received a maximum number of votes, discussions and voting continued until a primary customer was identified. The results were tabulated as follows. If there were five customers, then a first place vote would get five points, a second place vote would get four points, and so on. The customer with the largest number of votes was considered the primary customer. (Software Application used: Vote: Rank Order. Final time: 8 to 10 minutes)

**Step Four: Criteria Evaluation.** The purpose of Step Four was to evaluate each behavior on the final list against the critical attributes (i.e., customer satisfaction and process improvements) using an ordinal scale. By addressing the critical attributes, behaviors would be automatically assigned to a quadrant on the metric assessment tool. First, the facilitator asked the group to rate each behavior according to the following statement:

> As the primary customer, characterize your satisfaction with each behavior.

The group evaluated each behavior on a scale of "1" to "10". A "10" indicated that the customer would be extremely

satisfied with the behavior. A "1" indicated that the customer would be extremely dissatisfied or not affected by the behavior. The facilitator instructed the group members that they could also bypass their vote. If they elected to bypass, it would ensure discussions during the review of votes. The group members were also instructed that if they had trouble making a decision, they might want to consider the risk of the behavior to customer satisfaction. Once all votes had been entered (including bypasses), the facilitator reviewed the votes with the group. Any behavior with a bypass would be discussed. In addition, behaviors whose votes indicated a split of quadrants on the Metric Assessment Tool or a significant dispersion of results (e.g., three people vote "1" and two people vote "10") were also discussed. The purpose of the discussions was to ensure that all points of view were expressed. After discussions, group members were permitted to re-vote at their own option. The conclusion of the re-vote was considered final. Although some dispersion still existed, the mean allowed the computer to place the behavior on the metric assessment tool. If discussions were to continue for another round, the research team believes that they would have been forcing an unnatural response from the evaluation team. It was critical to capture the opinions of each group member without bias from the research team.

The facilitator then asked group members to rate each behavior in response to the following request:

> Determine the degree to which each behavior contributes to process improvement.

Once again group members evaluated each behavior on a scale of "1" to "10". A "10" indicated that the behavior definitely contributed to process improvement. A "1" indicated that the behavior definitely did not contribute to process improvement. Again, the facilitator instructed group members that they could bypass their vote. Upon the completion of the vote, discussions were held for bypass votes and significant dispersions to ensure that there was no confusion about the behavior or its impact to the process. The voting results were reviewed with the evaluation group and a re-vote was taken for those who wished to change their positions. The second vote was considered the final vote.

The final product of this step was a mean score indicating each behaviors' contribution to customer satisfaction and process improvement. This information was used to assign all behaviors to the appropriate quadrants of the Metric Assessment Tool. Behaviors were placed into quadrants automatically by the software according to their mean score. The mean is not the best measure of central tendency for an ordinal scale. However, the research team evaluated the median scores for both demonstrations and found no difference in the placement of behaviors based on

the median. Therefore, the research team relied upon the software (mean scores) to place behaviors into quadrants. The actual placement of behaviors by the software is demonstrated as follows.

| | Customer Satisfaction Mean Score | Process Improvement Mean Score |
|---|---|---|
| Quadrant A: | 5.6 to 10 | 5.6 to 10 |
| Quadrant B: | 5.6 to 10 | 1 to 5.5 |
| Quadrant C: | 1 to 5.5 | 5.6 to 10 |
| Quadrant D: | 1 to 5.5 | 1 to 5.5 |

Behaviors with a mean score from "5.6" to "10" on the customer satisfaction scale and a "5.6" to "10" on the process improvement scale were place in Quadrant A. Behaviors with a mean score from "1" to "5.5" on the customer satisfaction scale and "1" to "5.5" on the process improvement scale were placed in Quadrant D, etc. (Software Application used: Alternative Evaluator. Final time: 40 to 50 minutes)

**Step Five: The Metric Assessment Tool.** The purpose of Step Five was to provide the results of the evaluation in Step Four and to explain the significance of the metric assessment tool. Each group member was provided a metric assessment tool diagram which listed each behavior within its assigned quadrant (Appendix I for S&T group and Appendix J for B-2 group). In addition, the group was provided with a description of the tool which explained the significance of each quadrant (Appendix K). Each quadrant was explained to the group with examples from their own

evaluation and time was provided for questions and clarifications. The explanation included a discussion of the measurement deficiencies associated with certain quadrants. The final product of Step Five was used as guidance for evaluating the organizational objective in Step Six and for generating measurement improvement actions in Step Seven. (Final time: 30 minutes)

**Step Six: Evaluating the Objective.** The purpose of Step Six was to determine the degree to which a metric supports an organizational objective. As discussed in Chapter II, measurements should drive appropriate behaviors which support organizational objectives. The organizational objective associated with the metric under evaluation was provided by the office focal point. To conduct the evaluation, group were asked to address two questions and rate their responses on an ordinal scale. First, the facilitator asked the group to review the organizational objective and behaviors in all four quadrants as a composite. The facilitator asked the following question:

> How well do the composite behaviors support the organizational objective?

Each group member addressed this question on a scale of "1" to "10". A "10" indicated that the composite behaviors supported the organizational objective very well. A "1" indicated that the composite behaviors poorly supported the organizational objective. The facilitator then asked group

members to consider the behaviors in quadrant A only. The facilitator asked the following question:

> How well do the Quadrant A behaviors support the organizational objective?

Once again each group member addressed the question on a scale of "1" to "10". A "10" indicated that Quadrant A behaviors supported the organizational objective very well. A "1" indicated that Quadrant A behaviors poorly supported the organizational objective.

Each evaluation served a specific purpose. The first evaluation indicated the degree to which the metric supported the organizational objective prior to any recommendations for process measurement improvements. The second evaluation indicated whether the process measurement improvement actions to be identified in Step Seven, would enhance the support of the organizational objective. The purpose of the process measurement improvement actions in Phase I of step seven are to remove or control behaviors in Quadrants B, C, and D. Therefore, by using the behaviors in Quadrant A, the group was able to ascertain whether the behaviors driven by the metric (considered appropriate and satisfactory to the customer) would enhance the support of the objective. For Step Six the median scores were used. Median scores are considered the most appropriate measure of central tendency for ordinal scales. One median score was calculated for each evaluation question. If the second median score (Quadrant A median) was below the first median

score (composite median) then specific measurement improvement actions regarding the relationship between the metric and the objective would be solicited in Phase III in Step Seven. (Software Application used: Vote: Ten-Point Scale. Final time: 10 minutes)

**Step Seven: Improving the Process Measurement.** The purpose of Step Seven was to generate process measurement improvement actions. There were three phases to this step. During phase I the group provided recommendations for each deficient behavior (Quadrants B, C, and D) identified by the Metric Assessment Tool. During phase II the group provided recommendations to improve the metric structure itself. Finally, during phase III the group provided any other recommendations they had for improving the process measurement or its relationship to the organizational objective. Each phase was critical. The research team wanted to provide managers with maximum flexibility (options). Changes to a process measurement could take time and might be hampered by other constraints. Therefore, a manager might choose to implement some of the behavioral controls immediately while delaying other, more complicated, changes to a later time.

In phase I the evaluation group was asked to examine the behaviors in Quadrants B, C and D and generate recommendations for the process measurement deficiencies

discussed in Step Five. The recommendations were to be directed at eliminating or controlling the negative aspects of these behaviors. The following example was provided to the group.

| | |
|---|---|
| Behavior: | people would increase the amount of overtime to complete a project on time. |
| Deficiency: | Behavior would distort the true capabilities of the process and its measurement. |
| Possible Recommendations: | Management approval of overtime.<br>Work the process without overtime for a period to obtain accurate measurements. |

The evaluation group was instructed that not all behaviors may have a solution but that the intent of this exercise was to eliminate or control as many deficient behaviors as possible.

In phase II, the group as asked to generate improvement actions based on the structure of the metric. In other words, generate ideas to change the metric in an effort to eliminate undesirable behaviors. The group was given the following example:

> If a behavior indicates that quantity is emphasized over quality the group might suggest a different metric ratio, a supporting metric, or some other type of supporting measurement device such as a customer satisfaction survey.

All ideas were required to be as complete as possible.

In phase III, the group was asked to identify any other process measurement improvement actions which might not have been covered thus far. The facilitator told the group that they might want to consider the relationship between the organizational objective and the metric. In other words, what possible suggestions could the group make which would enhance this relationship?

The group was instructed that all recommendations in Step Seven should be unconstrained. If necessary, they could suggest a change in the metric, a change in the objective (or the creation of a subobjective), or any other actions which might improve the process measurement. In addition, the group was instructed to consider only the metric under evaluation and its objective.

The generation of possible solutions was done independently using the GSS. The product of this step was a list of proposed measurement improvement actions. (Software Application used: Categorizer. Final time 40 to 50 minutes)

**Step Eight: Post Discussions.** The purpose of Step Eight was to clarify the possible solutions identified in Step Seven to the research team through group interaction. After pretest II, the research team had decided to eliminate this step. However, after the first demonstration it became clear that some clarification was

necessary. Therefore, for the second demonstration the facilitator went through each recommendation to ensure that the research team had a correct understanding of the suggestion. The product of Step Eight was a list of measurement improvement actions which the group believes will improve the quality of the process measurement. This list was provided to the office focal points for possible implementation. (Final time: 15 minutes)

Final Comments. Upon the completion of the metric assessment, comments were solicited about the metric assessment tool and the metric assessment process. To construct the questionnaire the research team centered questions around the following three concepts:

(1) the content validity of the MAT,

(2) the understandability of the MAT and the assessment process, and

(3) the value of the MAT and the assessment process in evaluating metrics.

An open-response questionnaire was used so that participants could freely submit their thoughts without constraint or bias from the research team. The questionnaire was presented to the mock evaluation group in pretest II and responses indicated a satisfaction of the questionnaire concepts. The objective of the research effort was to develop and demonstrate a metric assessment tool. As such, the research team solicited comments and made improvements

to the tool or process upon the completion of each demonstration. A copy of the questionnaire provided to individuals in each evaluation group is included as Appendix L.

## Summary

Table 3.5 illustrates the mapping of the methodology to the research subobjectives posed in Chapter I.

**TABLE 3.5**

**RESEARCH METHODOLOGY SUMMARY**

| RESEARCH SUBOBJECTIVES | METHODOLOGY | PAGE # |
|---|---|---|
| Identify critical attributes | Literature Review | 3-2 |
| Develop metric assessment tool | Research Team | 3-2 |
| Determine if behaviors motivated by metric support the organization's goals and objectives | Group Evaluation | 3-33 |
| Identify behaviors which could distort the process measurement | Metric Assessment Tool/Group Evaluation | 3-5<br>3-24 |
| Determine the need for additional supporting metrics | Group Evaluation | 3-35 |
| Identify process measurement improvement actions | Group Evaluation | 3-35 |

# IV. Results

## Introduction

This chapter analyzes and reports the results of the metric assessments for the Science and Technology (S&T) Division and the B-2 System Program Office (SPO). The results of each demonstration are presented as received from the evaluation group. Therefore, corrections to grammar and syntax have not been made to any of the evaluation group data presented in table form throughout this chapter. The chapter is divided into three sections: the S&T demonstration, the B-2 SPO demonstration, and the results of the evaluation group questionnaires.

## Demonstration One: Science and Technology (S&T)

The following S&T metric met the selection criteria presented in Chapter Three:

> **S&T Teamwork Metric:** The percent of Advanced Technology Transition Demonstrations (ATTDs) with signed Technology Transition Plans (TTPs).

Definitions of an ATTD and a TTP are provided in Appendix A. The S&T Teamwork metric was developed to ensure the timely and accurate transition of technology by informing the S&T focal point, the product division, and the technology recipient (typically a SPO) of the criteria that must be met for the technology to be validated and accepted. These criteria are outlined in the TTP and every ATTD is expected

to have a signed TTP in place prior to the start date of the ATTD. The S&T Teamwork metric is identified with the following AFMC Command Objective:

> **Objective 3.2:** Transition technology rapidly to applications to include organic infrastructure.

Assessment Overview. As described in Chapter Three, the research team initially required five process knowledgeable individuals to act as the evaluation group. Due to personnel availability and time constraints, the actual S&T evaluation group consisted of only three individuals. Although the actual sample size was less than ideal, the researchers, in keeping with the research objective, concluded that three individuals would suffice to adequately demonstrate the metric assessment tool and process.

The instructions described in Chapter Three were presented to the evaluation group without exception. Each evaluation group member was provided an instruction package which contained: a Software Reference Guide (Appendix D), a process overview (Appendix E), and a Metric Assessment Fact Sheet (Appendix F). The S&T metric assessment demonstration contained seven steps. The results and analysis are now described.

Step One: Identification of Behaviors. In Step One, the evaluation group was asked the following question:

What behaviors might occur as a result of this metric?

Each evaluation group member anonymously entered behaviors into his/her computer terminal. The action was completed when new entries by the evaluation group had ceased. At this time, the facilitator asked the group if more time was required. All respondents indicated "No" with a thumbs down motion.

Step Two: Behavior Discussions and Refinements. The behavioral listing generated in Step One was refined during Step Two. The initial list contained a significant number of metric outcomes (as differentiated from metric behaviors in Chapter Three). Due to the nature of the metric assessment tool, behaviors are preferred because they typically identify the act and/or actor of a specific response. Therefore, during Step Two the facilitator attempted to instruct the group on the difference between outcomes and behaviors. The evaluation group then reviewed each entry. Changes were not made to the original list unless all evaluation group members agreed/disagreed with the modification, clarification or merger using a thumbs up/down voting system. Table 4.1 contains the modified S&T behavioral listing as it was received from the evaluation group. The group participants required extensive discussions when modifying the original list. Much of the groups' discussion centered around the meaning or intent of

**TABLE 4.1**

**S&T TEAMWORK METRIC: BEHAVIOR PREDICTIONS**

| BEHAVIOR # | BEHAVIOR |
|---|---|
| 1 | Personnel (all levels) emphasize quantity over quality |
| 2 | Personnel motivated to complete TTPs in most cases |
| 3 | Increased desire by TTP author to understand process |
| 4 | Metric brought process sluggishness to management attention |
| 5 | Two-letter management pays more attention to quality in process results |
| 6 | 2-letter management attention increases time to complete process |
| 7 | TTP signature level process is faster |
| 8 | Forced TTP author to handcarry document at times - work harder |
| 9 | Forced TTP author to interface/delegate duties - work smarter |
| 10 | Personnel emphasize quantity over process |
| 11 | Made process players more involved (i.e., boss is watching) |
| 12 | TTP author was often unaware of WL/XPT's schedule of events |
| 13 | Metric measured "one time" get well program |
| 14 | Personnel encouraged to interact with customer more often |
| 15 | Forced TTP schedules create encouragement to "get it done" |
| 16 | TTP author at Directorate level felt lack of authority over process |

the entries. For example, some group participants were unsure about the meaning of Behavior Number 11. Discussing this entry, the group focused on the following question: Does involvement speed up the process to the detriment of quality, or slow down the process to the customers' dismay? Similar discussions were held for every behavior listed. The modified list in Table 4.1 continues to contain outcomes (Behavior Numbers 4, 7, 11, 12, 13, 15, and 16). After discussions, the evaluation group elected not to modify the wording of these "behaviors".

Step Three: Voting on Primary Customer. The evaluation group was instructed to vote on the primary customer by placing the primary customer at the top of the list, followed by the others in decreasing order of importance. The primary customer was defined for the group as the customer who benefits the most, and the most directly, from the output of the process the metric measures. The potential customer list was obtained from the S&T focal point and pre-loaded into the computer software prior to beginning the assessment. The three customers identified were the developmental customers (typically a SPO), the Major Command or end user, and the Product Division (system developer and maintainer). An "Other" category was included to allow participants to identify additional customers not included in the prearranged list.

Table 4.2 contains the final results of the primary customer voting. All three

**TABLE 4.2**

**S&T TEAMWORK METRIC: RESULTS OF CUSTOMER VOTING**

| CUSTOMER | RANKING 1 | 2 | 3 | 4 | NUMBER VOTING |
|---|---|---|---|---|---|
| Developmental | 3 | - | - | - | 3 |
| Major Command | - | 2 | 1 | - | 3 |
| Product Division | - | 1 | 2 | - | 3 |
| Other | - | - | - | 3 | 3 |

evaluation group members ranked the developmental customer as the primary customer. Thus, the group felt the SPO was the benefactor of the output of the process being measured by the S&T Teamwork metric.

Step Four: Criteria Evaluation. During the criteria evaluation, participants were asked how well they felt each behavior contributed to customer satisfaction and process improvement. The evaluation group responded to the following computer directions:

> As the primary customer, characterize your satisfaction with each behavior,
>
> AND
>
> Determine the degree to which each behavior contributes to process improvement.

As described in Chapter Three, each behavior was scored using an ordinal scale from 1 to 10. A vote of "10"

indicated customer satisfaction and/or process improvement. A "1" vote indicated the customer may be dissatisfied or not affected by the behavior. The initial votes were cast and reviewed by the researchers. During the review, members of the evaluation group were able to see the voting results. Any split votes were discussed. A split vote occurred when some participants voted the mid-point value of 5.5 for any behavior. Discussions of the split votes revealed various interpretations of the impact or intent of the behavior. After discussing the split votes, group members were given the opportunity to cast new votes, if they desired, and final scores were tabulated. The mean responses were automatically calculated and are presented in Table 4.3. The actual behavior can be found by referencing Table 4.1.

Although the median is considered a better measure of central tendency, time constraints made the automatically tabulated mean more useable. Upon reexamination, the research team found that in every instance, the mean and the median both were either above or below the mid point value of 5.5

A mean value above 5.5 indicated that, on the average, the evaluation group felt the behaviors would result in customer satisfaction and/or process improvement. A mean score below 5.5 indicated that, on the average, the customer may be dissatisfied or not affected, and/or the process may

**TABLE 4.3**

**S&T TEAMWORK METRIC: CRITERIA EVALUATION RESULTS**

| BEHAVIOR # | CUSTOMER SATISFACTION MEAN RESPONSE | PROCESS IMPROVEMENT MEAN RESPONSE |
|---|---|---|
| 1 | 2.00 | 1.67 |
| 2 | 8.33 | 9.00 |
| 3 | 8.67 | 9.33 |
| 4 | 8.00 | 10.00 |
| 5 | 8.33 | 8.33 |
| 6 | 3.00 | 5.00 |
| 7 | 3.00 | 8.33 |
| 8 | 6.33 | 3.33 |
| 9 | 8.33 | 9.33 |
| 10 | 2.67 | 2.00 |
| 11 | 7.67 | 9.00 |
| 12 | 1.33 | 2.00 |
| 13 | 2.00 | 2.00 |
| 14 | 9.67 | 9.00 |
| 15 | 4.33 | 4.00 |
| 16 | 1.33 | 2.67 |

not be improved by the expected behavior. Mean or median values of 5.5 would indicate uncertainty among the evaluation group concerning a behavior's anticipated affect. Reexamination of the data revealed no mean or median values of 5.5. Table 4.3 indicates that the evaluation group essentially agreed on the impact of the behavior on the

process and the customer. Another result in Table 4.3 worth explanation reveals concerns Behavior Number 4 on process improvement. This behavior reads, "Metric brought process sluggishness to management attention." The behavior, by itself, does not indicate any specific management action. However, the evaluation group rated it as though this behavior would generate process improvements. Clearly, taken independently, management's attention is not an action that definitely improves upon a process. Thus, it appears that the evaluation group did not consider each behavior as an independent action, thereby reading more into the behavior than is apparent to the casual observer. Finally, Table 4.3 indicates that the evaluators voted toward the scale extremes. This is indicated by the numerous mean responses approaching "1" or "10" on the 10-point scale. This voting resulted in the majority of the behaviors being assigned to MAT Quadrants A and D during Step Five.

Step Five: The Metric Assessment Tool. A summary of the Metric Assessment Tool (MAT) for the S&T demonstration is included as Figure 4.1. Reference Table 4.1 for the actual behaviors. A MAT with a full description of all behaviors in included as Appendix I.

Behaviors were placed onto the MAT using the mean scores and the methodology described in Chapter Three. For example, Behavior Number 8 received a 6.33 customer

satisfaction score indicating the evaluation group expected the customer to be satisfied with the behavior. Likewise, the process improvement score of 3.33 indicated that the group felt the behavior would not improve the process. Thus, Behavior Number 8, "Working harder", landed in MAT Quadrant B. This indicates that working harder may satisfy a customer, but does not improve the process.

| | | |
|---|---|---|
| **CUSTOMER SATISFACTION** | **A**<br>Behavior No. 2<br>Behavior No. 3<br>Behavior No. 4<br>Behavior No. 5<br>Behavior No. 9<br>Behavior No. 11<br>Behavior No. 14 | **B**<br>Behavior No. 8 |
| **NO CUSTOMER SATISFACTION** | **C**<br>Behavior No. 7 | **D**<br>Behavior No. 1<br>Behavior No. 6<br>Behavior No. 10<br>Behavior No. 12<br>Behavior No. 13<br>Behavior No. 15<br>Behavior No. 16 |
| | **IMPROVEMENT ACTIONS** | **SUSTAINING ACTIONS** |

**Figure 4.1 The S&T Metric Assessment Tool**

Once behaviors were identified by quadrant, copies of the MAT were provided to the evaluation group and the significance of the behavior locations discussed in accordance with Appendix K.

Step Six: Evaluating the Objective. Using the behaviors as identified in the MAT (Appendix I), each group member was asked whether or not they felt the metric supported the organizational objective. The objective was provided on the Metric Assessment Fact Sheet (Appendix F). The evaluation was done in two parts as described in Chapter Three. First, the evaluation group members considered all behaviors found on the MAT as a composite. Next, participants considered only those behaviors determined to satisfy a customer and improve the process (those behaviors located in Quadrant A on the MAT). Again, an ordinal scale was used by the evaluation group. A "1" response indicated the behaviors generated by the metric poorly supported the organizational objective. A "10" indicated the behaviors generated by the metric supported the objective very well. The median response figures are presented as follows in Table 4.4.

**TABLE 4.4**

**S&T TEAMWORK METRIC: OBJECTIVE VOTING RESULTS**

| BEHAVIORS | MEDIAN SCORE |
|---|---|
| Composite | 6.00 |
| Quadrant A Only | 7.00 |

Although there was not a great deal of difference between the two median scores, the difference does appear to indicate that if deficient behaviors (behaviors in Quadrants

B, C, and D) were eliminated, the metric would support the objective to a somewhat greater degree. This did not indicate whether or not the objective itself was good or bad, only the degree to which behaviors motivated by the metric supported the objective. Thus, the evaluation group determined that deficient behaviors hinder the accomplishment of the objective.

Step Seven: Improving the Process Measurement. Three categories of improvement were considered during Step Seven as identified in Chapter Three. The improvement actions are presented below. No effort has been made to alter or correct the grammar or syntax of any recommendations.

In Phase I, the evaluation group focused on controlling or eliminating deficient behaviors (behaviors in Quadrants B, C, and D). Table 4.5 lists the deficient behavior and the associated improvement recommendation(s) as identified by the evaluation group.

**TABLE 4.5**

**S&T TEAMWORK METRIC: INDIVIDUAL BEHAVIOR RECOMMENDATIONS**

| **IMPROVEMENT ACTIONS: INDIVIDUAL BEHAVIORS** |
|---|
| **BEHAVIOR (Behavior No.)**<br>• **IMPROVEMENT ACTION(S)** |
| Personnel (all levels) emphasize quantity over quality (1)<br>• Stress "back to basics" theory to improve quality<br>• Develop training that stresses value of quality TTP |

**TABLE 4.5**

**S&T TEAMWORK METRIC: INDIVIDUAL BEHAVIOR RECOMMENDATIONS (CONTINUED)**

| **IMPROVEMENT ACTIONS: INDIVIDUAL BEHAVIORS** |
|---|
| **BEHAVIOR (Behavior No.)**<br>• **IMPROVEMENT ACTION(S)** |
| 2-letter management attention increases time to complete process (6)<br>• Educate 2-letters on process and their part in it<br>• Decrease time through short high-level briefings |
| TTP signature level process is faster (7)<br>• Eliminate requirement for "formality" signature |
| Forced TTP author to handcarry doc. - work harder (8)<br>• Eliminate need for "hand carried" TTP's<br>• Develop reasonable timeline guidance<br>• Standardize timeliness - for all TTP actions - be consistent<br>• "Fix" the system to avoid crisis management |
| Personnel emphasize quantity over process (10)<br>• Educate personnel with a "fixed" TTP process guidance<br>• Provide meaningful/consistent one-time guidance to directorate |
| TTP author often unaware of WL/XPT sched. of events (12)<br>• Give author more responsibility for scheduling<br>• Educate personnel with a "fixed" TTP process guidance |
| Metric measured "one time" get well program (13)<br>• Develop customer satisfaction metric<br>• Explore "all options for other candidate metrics |
| Forced TTP schedules encourage "getting it done" (15)<br>• Offer more opportunity for TTP author involvement with schedule |
| TTP author at directorate level felt lack of authority over process (16)<br>• All directorates to modify TTP process<br>• Have scheduled status meetings between WL/XPT and Directorate |

In Phase II, the evaluation group was asked to consider improvements to the Teamwork metric structure. Table 4.6 lists the improvements identified by the evaluation group for the metric structure.

In Phase III, the evaluation group was given the opportunity to include any additional recommendations for improving the measurement of the process. Table 4.7 contains the additional recommendations. This step completed the assessment process for the first demonstration.

**TABLE 4.6**

**S&T TEAMWORK METRIC: STRUCTURE IMPROVEMENTS**

| IMPROVEMENT ACTIONS: STRUCTURE |
|---|
| • Change TTP signature requirement date to be in line with schedule |
| • Add customer satisfaction metrics; Is customer happy? |
| • Do not single out specific program or laboratory |
| • Eliminate the current "headcount" metric - use spreadsheets |
| • Eliminate confusing guidance on TTPs and Metrics |
| • Need to measure whether SPOs are actually using developed technology. ie need to measure the effectiveness of the process by whether or not SPOs ever use technology. TTP's are only the plan for transition |
| • Publish metric - TTP tools for all involved in process |
| • Determine value added - To directors, Lab, SPOs, Customers |
| • Ask TTP authors if process works |

**TABLE 4.7**

**S&T TEAMWORK METRIC: ADDITIONAL IMPROVEMENTS**

| IMPROVEMENT ACTIONS: OTHER |
|---|
| • Establish TTP checklist |
| • Review checklists at Directorate level (may have WL/XPT presence) |
| • Establish action items to "fix" trouble spots |
| • Maintain customer reviews - use offsite style? Narrator? |
| • Provide "value added" feedback to management - including AFMC |
| • Decentralize control of laboratory process. Allow each directorate to develop their own process within a set of standards. This would allow for experimentation with better processes and benchmarking of the best process. |
| • In order to improve the metrics process, management must be aware of what value any changes may bring to bear on the issue |

Lessons Learned. The objective of this research effort was to develop and demonstrate a Metric Assessment Tool. In keeping with that objective, the research team identified the following problems associated with the metric assessment process.

1. The difference between outcomes and behaviors required further clarification.

2. The connotation of certain behaviors required a group consensus prior to Step Four to ensure evaluators retained knowledge concerning the meaning or intent of each behavior.

3. The research team required additional clarification concerning the improvement recommendations identified in Step Seven.

Changes to the assessment process were made to accommodate these concerns prior to the second demonstration.

## Demonstration Two: The B-2 Systems Program Office (SPO)

The following B-2 metric met the selection criteria presented in Chapter Three:

> **B-2 Responsiveness Metric:** The number of days between the receipt of an Engineering Change Proposal or Contract Change Proposal (ECP/CCP) and the recommendation for action from the Configuration Control Board (CCB).

Weapon system changes are submitted by private contractors as an ECP or a CCP. These changes are initially reviewed by entry level personnel in the B-2 SPO and forwarded to the CCB. The CCB either approves or disapproves the change. If the change is approved, the change is incorporated into the contract. The metric measures the number of days it takes for a CCB decision (either positive or negative) to be made. The B-2 metric has been named the "Responsiveness" metric by the researchers in an effort to ease differentiation between the two demonstrations.

The metric assessment process described in Chapter Three required each metric to be associated with an objective. This was necessary for evaluation purposes in accordance with the critical attributes listed in Chapter Two and in Step Six of the assessment process.

Additionally, the research sponsor rquired each metric demonstration to be logically linked to an AFMC Command Objective. The B-2 focal point revealed that the metric existed without an explicit objective or logical link to a command objective. Therefore, the B-2 Responsiveness Metric was identified with the following implicit objective (used to evaluate the metric's support of the organizational objective during Step Six of the assessment process):

> **Objective:** Minimize the time it takes to incorporate an idea (for product enhancement or deficiency correction) into a contract.

Additionally, the Responsiveness Metric was logically linked to AFMC Command Objective 1.1 by the researchers and the B-2 focal point.

> **Objective 1.1:** Understand, through sustained interaction, our customers and their requirements, and provide options, including those available through other services, which are the basis for customer decisions and satisfaction.

Assessment Overview. As described in Chapter Three, the research team initially required five process knowledgeable individuals to act as the evaluation group. Due to personnel availability and time constraints, the actual B-2 evaluation group consisted of only four individuals. Although the actual sample size was less than ideal, the researchers, in keeping with the research objective, concluded that four individuals would suffice to

adequately demonstrate the metric assessment tool and process.

The lessons learned during the first assessment were addressed prior to this demonstration and incorporated into the assessment process. All instructions described in Chapter Three were presented to the evaluation group without exception. Each evaluation group member was provided an instruction package which contained: a Software Reference Guide (Appendix D), a process overview (Appendix E), a Metric Assessment Fact Sheet (Appendix G), and a metric outcome and connotation example (Appendix H). The B-2 metric assessment demonstration contained eight steps. The results and analysis are described below.

Step One: Identification of Behaviors. In Step One, the evaluation group was asked the following question:

What behaviors might occur as a result of this metric?

This question was preceded by a discussion concerning the difference between outcomes and behaviors (as described in Chapter Three and in Appendix H). This action was in response to the confusion identified during the first demonstration. The group members anonymously entered behaviors into their terminals. The action was completed when new entries by the evaluation group had ceased. At this time, the facilitator asked the group if more time was

required. All respondents indicated "No" with a thumbs down motion.

Step Two: Behavior Discussions and Refinements. The behavioral listing generated in Step One was refined during Step Two. Unlike the S&T demonstration, the initial list contained metric behaviors, not outcomes (this resulted in less overall refinement time). The first demonstration indicated a need for each member to be fully aware of a behavior's connotation. Therefore, prior to reviewing each behavior generated in Step One, the facilitator provided the group with an "it depends" example (Appendix H). A behavior that could be construed as either desirable or undesirable, depending on the situation, was divided into its positive and negative components. Behaviors were assigned a plus symbol (+) if they were desirable, a minus symbol (-) if they were undesirable, or no symbol if the behaviors connotation was easily derived. Modifications, clarifications and mergers were completed by the group and all changes were voted on using a thumbs up/down voting system. Table 4.8 contains the modified behavioral listing for the responsiveness metric and is presented, as received, from the evaluation group without correction to grammar or syntax.

Unlike the first demonstration (S&T), the B-2 behavioral listing was less ambiguous. Almost every

**TABLE 4.8**

**B-2 RESPONSIVENESS METRIC: BEHAVIOR PREDICTIONS**

| BEHAVIOR # | BEHAVIOR |
|---|---|
| 1 | Effort expended to get around the process instead of using it correctly |
| 2 | Encourages putting the measured proposal as the top priority (+) |
| 3 | Could cause one to neglect other work |
| 4 | People concerned with due dates rather than quality technical proposals |
| 5 | Managers may use the metric as a personnel evaluation tool (-) |
| 6 | Would try to do more up front coordination to facilitate process |
| 7 | could cause you to seek proposal review comments before suspense due (+) |
| 8 | Might cause one to schedule CCB date prematurely |
| 9 | Try to do "cya" activities, like blame the contractor |
| 10 | May cause "unhealthy competition" between teams |
| 11 | May increase team motivation/cooperation/ creativity |
| 12 | May induce conflict between process owners and process users |
| 13 | Try to have proposal designated as the number one priority in the SPO (-) |
| 14 | Would lead to creative ways to define requirements and deal with the KR |
| 15 | Lead to inappropriate actions in dealing with contractors |

**TABLE 4.8**

**B-2 RESPONSIVENESS METRIC: BEHAVIOR PREDICTIONS (CONTINUED)**

| BEHAVIOR # | BEHAVIOR |
|---|---|
| 16 | Could cause one to ask contractor for proposal change pages (+) |
| 17 | May cause unnecessary TDY to correct deficiencies with the contractor |
| 18 | Could cause extended work hours (working harder) |
| 19 | Could cause unnecessary TDY to correct deficiencies with the contractor |
| 20 | Encourages modification of work habits/ increase efficiency |

behavior had a specific act and/or actor associated with it. Additionally, the discussion on each behavior connotation helped to clear up any confusion concerning the behavior's meaning or intent.

Step Three: Voting on Primary Customer. The evaluation group was instructed to vote on the primary customer by placing the primary customer at the top of the list, followed by the others according to their importance. The primary customer was defined as the customer who benefits the most, and the most directly, from the output of the process the metric measures. The potential customer list was obtained from the B-2 focal point and pre-loaded into the computer software prior to beginning the

assessment. The three customers identified were the B-2 SPO, the Major Command (Air Combat Command), and the primary contractor for the B-2 (Northrop Corporation). An "Other" category was included to allow participants to identify additional customers not included in the prearranged list.

The evaluation group was initially split between whether the Major Command or the SPO was the primary customer (half of the members voted for each). The facilitator allowed the group time to discuss the reasons behind each vote. After discussing the definition given by the facilitator of the primary customer, the group voted a second time. The results of the final voting are found in Table 4.9.

The B-2 SPO was considered the primary customer of the output of the process by receiving all four first place votes.

**TABLE 4.9**

**B-2 RESPONSIVENESS METRIC: RESULTS OF CUSTOMER VOTING**

| CUSTOMER | RANKING 1 | 2 | 3 | 4 | NUMBER VOTING |
|---|---|---|---|---|---|
| SPO | 4 | - | - | - | 4 |
| Major Command | - | 4 | - | - | 4 |
| Contractor | - | - | 3 | 1 | 4 |
| Other | - | - | 1 | 3 | 4 |

Step Four: Criteria Evaluation. During the criteria evaluation, participants were asked how well they felt each behavior contributed to customer satisfaction and process improvement. The evaluation group responded to the following computer directions:

> As the primary customer, characterize your satisfaction with each behavior,
>
> AND
>
> Determine the degree to which each behavior contributes to process improvement.

As described in Chapter Three, each behavior was scored using an ordinal scale from 1 to 10. A vote of "10" indicated customer satisfaction and/or process improvement. A "1" vote indicated the customer may be dissatisfied or not affected by the behavior. The initial votes were cast and reviewed by the researchers. During the review, members of the evaluation group were able to see the voting results. Any split votes were discussed. A split vote occurred when participants voted both above and below the mid-point value of 5.5 for any behavior. After discussing the split votes, group members were given the opportunity to cast new votes, if they desired, and final scores were tabulated. The mean responses are presented in Table 4.10. The actual behavior can be found by referencing Table 4.8.

Although the median is considered a better measure of central tendency, time constraints made the automatically tabulated mean more useable. Upon reexamination, the

**TABLE 4.10**

**B-2 RESPONSIVENESS METRIC: CRITERIA EVALUATION RESULTS**

| BEHAVIOR # | CUSTOMER SATISFACTION MEAN RESPONSE | PROCESS IMPROVEMENT MEAN RESPONSE |
|---|---|---|
| 1 | 1.00 | 1.75 |
| 2 | 6.00 | 2.25 |
| 3 | 2.75 | 2.00 |
| 4 | 2.00 | 3.25 |
| 5 | 3.00 | 1.50 |
| 6 | 9.00 | 8.75 |
| 7 | 5.75 | 7.25 |
| 8 | 3.25 | 3.75 |
| 9 | 2.50 | 1.75 |
| 10 | 2.50 | 1.50 |
| 11 | 8.75 | 8.25 |
| 12 | 2.25 | 1.50 |
| 13 | 2.26 | 1.50 |
| 14 | 8.25 | 9.00 |
| 15 | 1.00 | 1.00 |
| 16 | 8.00 | 8.50 |
| 17 | 1.00 | 1.25 |
| 18 | 2.50 | 1.75 |
| 19 | 2.00 | 1.75 |
| 20 | 8.50 | 9.25 |

research team found that in every instance, except one, the mean and the median placed the behavior in the same quadrant on the MAT. The one exception, Behavior Number 7, had a median value of 5.5 for customer satisfaction. A mean value above 5.5 indicated that the evaluation group members felt that the behavior would result in customer satisfaction and/or process improvement. A mean score below 5.5 indicated the customer may be dissatisfied or not affected, and/or the process may not be improved by the expected behavior. Mean or median values of 5.5 indicated uncertainty among the evaluation group concerning a behavior's anticipated affect. Reviewing the means and medians revealed that the evaluation group felt each behavior was either a positive or negative influence upon the customer and/or process. Behavior Number 7, however, was the exception to this rule.

Behavior Number 7 reads, "Could cause you to seek proposal review comments before suspense due...(+)." Two of the group evaluators rated this behavior a "5" on customer satisfaction while the other two rated it a "6" and "7" respectively, making the mean 5.75 and the median 5.5. Since the median value was neither above nor below the mid-point, the research team concluded that the mean value was still an acceptable indicator of the central tendency. This problem may have been avoided had the evaluation group size been an odd number.

A split vote occurred on Behavior Number 2 which stated, "Encourages putting the measured proposal as the top priority (+)." Individual voting revealed two participants felt this behavior was a customer satisfier. scoring it an "8", while the other participants scored the behavior a "5" and a "3" respectively, on the 10-point scale. This may be indicative of different customer perspectives. If the B-2 SPO proposal was given the top priority, the B-2 SPO would be satisfied. However, if the B-2 SPO proposal was not given top priority, they would be dissatisfied.

Finally, Table 4.10 reflects the tendency of evaluators to vote toward the scale extremes. This is indicated by the numerous mean responses approaching "1" or "10" on the 10 point scale. This voting resulted in the majority of behaviors being assigned to MAT Quadrants A and D during Step Five. A similar observation was made concerning the results obtained during the S&T demonstration.

Step Five: The Metric Assessment Tool. A summary of the Metric Assessment Tool (MAT) for the B-2 demonstration is included as Figure 4.2. (reference Table 4.8 for the actual behaviors). A MAT with a full description of all behaviors in included as Appendix J. Behaviors were placed onto the MAT using the mean scores and the methodology described in Chapter Three. For example, Behavior Number 2 received a 6.00 customer satisfaction score indicating the

| | A | B |
|---|---|---|
| CUSTOMER SATISFACTION | Behavior No. 6<br>Behavior No. 7<br>Behavior No. 11<br>Behavior No. 14<br>Behavior No. 16<br>Behavior No. 20 | Behavior No. 2 |
| NO CUSTOMER SATISFACTION | C | D<br>Behavior No. 1<br>Behavior No. 3<br>Behavior No. 4<br>Behavior No. 5<br>Behavior No. 8<br>Behavior No. 9<br>Behavior No. 10<br>Behavior No. 12<br>Behavior No. 13<br>Behavior No. 15<br>Behavior No. 17<br>Behavior No. 18<br>Behavior No. 19 |
| | IMPROVEMENT ACTIONS | SUSTAINING ACTIONS |

**Figure 4.2 The B-2 Metric Assessment Tool**

evaluation group expected the customer to be satisfied with the behavior. Likewise, the process improvement score of 2.25 indicated that the group felt the behavior would not improve the process. Thus, Behavior Number 2 landed in MAT Quadrant B. This indicates that the behavior satisfies the customer but is only sustaining, not improving, the process. Once behaviors were identified to quadrants, copies of the

MAT were provided to the evaluation group and the significance of the behavior locations discussed in accordance with Appendix K.

The MAT revealed a critical research result. Specifically, "working harder" behaviors, such as Behavior Number 18, were expected to fall within Quadrant B. Though working harder does not improve a process, the researchers felt that working harder meant working harder for the customer who would become satisfied by the extra effort. However, in this instance, "Working harder" was associated with "...extended work hours" which the evaluation group felt would dissatisfy, as differentiated from "not affecting", the B-2 SPO primary customer. This distinction is made since the B-2 SPO customer is directly affected by the process workload within its organization. Increased overtime would dissatisfy the B-2 management, as predicted by the evaluation group, since the process is unable to sustain the demand placed upon it. However, from another customer's perspective (the Major Command), working harder to meet product requirement, with or without overtime, may be satisfying.

Step Six: Evaluating the Objective. Using the behaviors as identified in the MAT (Appendix J), each group member was asked whether or not they felt the metric supported the organizational objective. The objective was

provided on the Metric Assessment Fact Sheet (Appendix G). The evaluation was done in two parts as described in Chapter Three. First, the evaluation group members considered all behaviors found on the MAT as a composite. Next, participants considered only those behaviors determined to satisfy a customer and improve the process (those behaviors located in Quadrant A on the MAT). Again, an ordinal scale was used by the evaluation group. A "1" response indicated the behaviors generated by the metric poorly supported the organizational objective. A "10" indicated the behaviors generated by the metric supported the objective very well.

Although there was not a great deal of difference between the two median scores, the difference does appear to indicate that if deficient behaviors (behaviors in Quadrants B, C, and D) were eliminated, the metric would support the objective to a slightly greater degree. This did not indicate whether or not the objective itself was good or bad, only the degree to which the behaviors motivated by the metric supported the objective. Thus, the evaluation group determined that deficient behaviors hinder the accomplishment of the objective. The median response figures are presented as follows in Table 4.11.

**TABLE 4.11**

**B-2 RESPONSIVENESS METRIC: OBJECTIVE VOTING RESULTS**

| BEHAVIORS | MEDIAN SCORE |
|---|---|
| Composite | 6.00 |
| Quadrant A Only | 6.50 |

Step Seven: Improving the Process Measurement. As identified in Chapter Three, three categories of improvement were considered during Step Seven. The improvement actions are presented below. No effort has been made to alter or correct the grammar or syntax of any recommendations.

In Phase I, the evaluation group focused on controlling or eliminating deficient behaviors (behaviors in Quadrants B, C, and D). Table 4.12 lists the deficient behavior and the associated improvement recommendation(s) as identified by the evaluation group.

In Phase II, the evaluation group was asked to consider improvements to the Teamwork metric structure. Table 4.13 lists the improvements identified by the evaluation group for the metric structure.

In Phase III, the evaluation group was given the opportunity to include any additional recommendations for improving the measurement of the process. Table 4.14 contains the additional recommendations.

**TABLE 4.12**

**B-2 RESPONSIVENESS METRIC: INDIVIDUAL BEHAVIOR RECOMMENDATIONS**

| IMPROVEMENT ACTIONS: INDIVIDUAL BEHAVIORS |
|---|
| **BEHAVIOR (Behavior No.)**<br>• **IMPROVEMENT ACTION(S)** |
| Effort expended to get around the process instead of using it correctly (1)<br>• Have a clear policy on adherence to appropriate procedures |
| Effort expended to get around the process instead of using it correctly (1) (Continued)<br>• Establish ground rules regarding acceptable process practices<br>• Provide training to emphasize the importance of following the process<br>• Provide training to personnel using the process |
| Encourages putting the measured proposal as the top priority (2)<br>• Management must present consistent policy on workload<br>• Encourage users to maintain work priority lists on computer |
| Could cause one to neglect other work (3)<br>• Coordinate work priority with supervisor<br>• Periodic review of pending actions with supervisor<br>• When over worked, encourage employees to talk with manager<br>• Workers and managers work together to define a workers priorities |
| People concerned with due dates rather than quality technical proposals (4)<br>• Ensure metrics are used to evaluate process not performance<br>• Emphasize quality versus schedule |
| Managers may use the metric as a personnel evaluation tool (5)<br>• Emphasize to employees that quality of work is primary goal<br>• Do not use metric for personnel evaluations<br>• Managers should ensure evaluations are based on quality of effort |

**TABLE 4.12**

**B-2 RESPONSIVENESS METRIC: INDIVIDUAL BEHAVIOR RECOMMENDATIONS (CONTINUED)**

| IMPROVEMENT ACTIONS: INDIVIDUAL BEHAVIORS |
|---|
| **BEHAVIOR (Behavior No.)**<br>• **IMPROVEMENT ACTION(S)** |
| Managers may use the metric as a personnel evaluation tool (5) (Continued)<br>• Publish metric values by organization, not individual |
| Might cause one to schedule CCB date prematurely (8)<br>• CCB policy for explanation to board on why schedule change<br>• Require that all comments be resolved prior to CCB scheduling |
| Try to do "cya" activities, like blame contractor (9)<br>• Encourage documentation of effort so process improvements can be suggested<br>• Gather all the facts before laying blame<br>• Keep necessary parties informed of actions/decisions<br>• Document day-to-day work decisions |
| May cause "unhealthy competition" between teams (10)<br>• Encourage exchange of information in professional atmosphere<br>• Listen to employees/team members concerns |
| May induce conflict between process owners and process users (12)<br>• Clear definition of responsibilities alleviate misunderstandings<br>• Encourage comments from users on process improvements |
| People will try to have their proposal designated as the number one priority in the SPO (13)<br>• Management must recognize this motivation in their employees and stop it |
| Lead to inappropriate actions in dealing with contractors (15)<br>• Encourage all proposal managers to coordinate on correspondence activities<br>• Enforce policy with reprimand appropriate to action |

**TABLE 4.12**

**B-2 RESPONSIVENESS METRIC: INDIVIDUAL BEHAVIOR RECOMMENDATIONS (CONTINUED)**

| IMPROVEMENT ACTIONS: INDIVIDUAL BEHAVIORS |
|---|
| **BEHAVIOR (Behavior No.)**<br>• **IMPROVEMENT ACTION(S)** |
| Lead to inappropriate actions in dealing with contractors (15) (Continued)<br>• Have all letters to KTR reviewed by contracting officer |
| May cause people to "fudge" numbers to make metric look good (17)<br>• Encourage honesty. Set good example by concentrating on the process<br>• Exposure of fudged numbers would embarrass and prohibit future fudging |
| Could cause extended work hours (working harder) (18)<br>• Supervisors must evenly distribute workload to avoid overtime<br>• Encourage supervisor awareness of worker workload and hours<br>• Have team members work together during peak periods |
| Could cause unnecessary TDY to correct deficiencies with the contractor (19)<br>• Use VTC instead of TDY<br>• Interaction not a bad thing, but all unnecessary travel should be avoided. Supervisors must control with fair policies |

**TABLE 4.13**

**B-2 RESPONSIVENESS METRIC: STRUCTURE IMPROVEMENTS**

| IMPROVEMENT ACTIONS: STRUCTURE |
|---|
| • Examine the sub-processes within the process |
| • Ensure contractor/SPO place equal emphasis on metric |

**TABLE 4.13**

**B-2 RESPONSIVENESS METRIC: STRUCTURE IMPROVEMENTS (CONTINUED)**

| IMPROVEMENT ACTIONS: STRUCTURE |
|---|
| • Quality on how many changes and revisions are received on proposals<br>• Measure how many times a proposal is scheduled for the board and is deferred by the board because of insufficient data to influence a board decision. This measurement would tend to indicate meeting schedule rather than quality of work<br>• Look at CCB itself, and the type of issues which are discussed<br>• Look at contractor's approach to providing proposals<br>• Metric may be looked at by management to assess how well activity workload is being distributed throughout the SPO<br>• Evaluate time lost due to TDYs, absences, lack of OPRs |

**TABLE 4.14**

**B-2 RESPONSIVENESS METRIC: ADDITIONAL IMPROVEMENTS**

| IMPROVEMENT ACTIONS: OTHER |
|---|
| • Have periodic pulsing of the process users to get a feel how the process works for them - survey<br>• Transmit proposals to the SPO electronically - Duplicate in-house<br>• Provide meaningful evaluation of what the metric tells you about the process to the users so they can understand why the measurement is important<br>• Distribute proposals electronically as WP files or images |

Step Eight: Post Discussions. The S&T demonstration revealed a need for additional clarification concerning the process measurement improvement actions. As such, the final step in the B-2 demonstration was a discussion and clarification of selected improvement actions as described in Chapter Three. Tables 4.12, 4.13, and 4.14 contain the clarified improvement actions.

Lessons Learned. The objective of this research effort was to develop and demonstrate a Metric Assessment Tool. In keeping with that objective, the research team identified several problems associated with the second demonstration of the assessment process.

1. Clarification of the primary customer (output or product) is necessary, and
2. the evaluation group size should be an odd number to avoid split voting.

**Questionnaire Results**

At the completion of each metric assessment, comments were solicited about the Metric Assessment Tool and the metric assessment process. The questions, located in Appendix L, centered around the following three concepts:

1. the content validity of the MAT,
2. the understandability of the MAT and the assessment process, and
3. the value of the MAT and the assessment process in evaluating metrics.

Summaries of the questionnaires are presented as Appendix M for the S&T demonstration and Appendix N for the B-2 demonstration.

Content Validity. The MAT's content validity was supported by the evaluation group. Each member was asked how many behaviors they had witnessed within the process. Five of the seven users had witnessed "Many" of the behaviors occurring within the process while the other two had personally witnessed "All" of the behaviors. With respect to covering the metric assessment issue, one respondent wrote, "Group dynamics was good - got to see the "other side" of the question. The rescoring after discussion was extremely good - will help your quality." In fact, all evaluation group participants recommended the MAT to others assessing metrics. One respondent stated that the MAT "Provides the discipline and procedure to effectively evaluate a metric."

Understandability. The evaluation group also responded to questions concerning the understandability of the MAT and the assessment process. All S&T evaluators felt the MAT was easy to understand; however, two of the three evaluators had difficulty answering the questions regarding customer satisfaction and process improvement. One respondent wrote,

"It was often difficult to establish relationship to customer or process."

Responding to this feedback, several changes were incorporated into the assessment process prior to the second demonstration. In the second demonstration, the B-2 evaluation group, as a whole, felt the MAT was easy to understand with one respondent stating, "Well defined and explained." None of the B-2 participants had difficulty answering the questions regarding customer satisfaction, process improvement, and/or objective support. One respondent stated that the assessment process was "painless to follow."

Lastly, all participants in the demonstrations felt the assessment process was aided significantly by the computer software, GroupSystems V, located at the Armstrong Laboratory.

Assessment Value. The evaluation group members responded to several questions aimed at determining the value of the assessment tool and process. All respondents personally found the MAT to be useful indicating the MAT "...reveals what is/is not satisfying to the customer" while providing "...a forum for metric assessment." In fact, every user would recommend the MAT to others wishing to evaluate metrics. The users felt the MAT was an "Excellent

tool to determine if (the) metric is meaningful to anyone but you and will achieve goals."

The evaluation group members also commented concerning the value of the improvement actions generated. Six of the seven users felt the improvement actions would improve upon the current approach to measuring the process. Every respondent, however, questioned management's resolve to adopt constructive changes. Responding to this question, one respondent summed it up by stating the improvements would enhance the current approach only "If adopted."

**Summary**

Chapter Four describes the results of the Metric Assessment Tool demonstrations as they actually occurred. The following chapter will provide the research team's conclusions regarding the demonstrations and will provide some recommendations for further research into the area of metrics.

# V. Conclusions and Recommendations

## Introduction

This chapter highlights the conclusions drawn by the researchers as a result of this study. The chapter also contains recommendations for follow on studies.

## Conclusions

The following nine conclusions were formed by the research team as a result of this thesis effort.

1. Metrics can motivate inappropriate behaviors.
2. The Metric Assessment Tool (MAT) and assessment process are valuable management techniques.
3. There is no such thing as a perfect metric.
4. Metrics that focus on the end results of a process (such as quotas) will outlive their usefulness and will not promote continuous process improvements.
5. The assessment process generates extreme behaviors.
6. The assessment process is not final.
7. Metrics should be born of objectives.
8. Identifying the customer is not easy.
9. GroupSystems Software V is a useful forum for group discussions and decisions, but the software is not mandatory.

Metric Motivation. Driving appropriate actions is one of the attributes of a good metric. To be appropriate, the action, or behavior, must satisfy the customer and lead to

process improvements. This thesis study demonstrated that metrics drive both appropriate and inappropriate actions. Some of these inappropriate actions distort the process measurement and give the user the impression the process is performing better than it actually is. These behaviors will either satisfy the customer without leading to process improvements (i.e., fighting fires) or will generate a process improvement that will not ultimately benefit the customer. The remaining inappropriate behaviors neither satisfy the customer nor lead to process improvements. The result of the distortion is that managers can miss improvement opportunities which might otherwise be apparent.

The evaluation group questionnaires indicated that "Many" to "All" of the metric behaviors identified have actually occurred within the process. The responses indicated that a "get it done on time" mentality existed within their organizations. In other words, the participants believed that a favorable metric measurement was more important than the possible learning to be achieved from a valid metric measurement. The research team believes that this attitude drives inappropriate behaviors which are not conducive to continuous improvements. Managers must ensure that everyone clearly understands the purpose of a metric, and how individual behaviors will, or will not, contribute to the organizations objectives.

The Value of the MAT/Assessment Process. The MAT and the assessment process generate realistic, usable, and valuable information to managers. Realism is supported by the fact that the process identifies actual behaviors, evaluates behaviors according to the critical attributes of a valid metric, and generates improvement actions from the actual people whose behavior must be controlled.

The MAT and the assessment process are useful because they enable people to recognize the connection between the metric and their own behaviors. The MAT identifies deficient behaviors and the process solicits recommendations for controlling (or eliminating) those behaviors. Because the recommendations are generated by the workforce personnel, they will be more willing to accept and adopt the controls without reluctance. In other words, the tool and process provide managers with insight into potential problems and ideas for dealing with those problems which have been approved by their workforce. Managers relying upon metric data must be aware of distorting behaviors which have an impact on the accuracy of their measurements. Distorted information might result in poor or inaccurate decisions.

Finally, the MAT and the assessment process are valuable. They provide a means to bridge the gaps that inherently exist between customers, workers, processes, and objectives. Information aimed at filling these gaps is

critical to any organization since dissatisfied customers, chaotic processes, or unmet objectives generally result in management and/or organizational failure.

The Perfect Metric. The research suggests all metrics have some potential for inappropriate behavior. The key is to continually motivate appropriate behaviors while controlling or eliminating inappropriate behaviors. As discussed earlier, the most important ingredient to accurate metric measurements is a clear understanding, by all individuals involved, of the metrics purpose. Metrics should not be used as a personnel performance tool. They are a process measurement tool and should be used to improve the process, not used to berate the workers within the process. Removing the fear associated with "poor" measurement figures is difficult but is enhanced by the manager's continuous and proper use of metrics. The deficient behaviors identified by the MAT and the assessment process provide a warning sign to managers that their employees may need further delineation of the metrics purpose. In many cases, this might be enough to circumvent the use of inappropriate behaviors. In other cases, the improvement actions generated from the assessment process will help the manager deter the exhibition of inappropriate behaviors.

Quota Metrics. Metrics which focus on the end results of a process (such as quotas) will outlive their usefulness and will not promote continuous process improvements. Metrics which measure the percent of something, similar to the S&T metric which measures the percentage of ATTDs with signed TTPs, have a potential to outlive their effectiveness for continuous improvement. For example, what would motivate continuous improvement if the metric indicates 100% effectiveness? Is there no more room for improvement? Must we keep measuring? Clearly, the S&T metric does not indicate the **quality** of the signed TTPs. Therefore, there might still be room for improvement, but without a proper measurement, no improvements will be sought. The S&T metric, by focussing on the end result of a process, is a quota. Such metrics do not support continuous improvement. With these types of metrics, the workforce is motivated to achieve the goal, the 100% effectiveness goal. Their motivation is with the measurement figure instead of with process improvement. Once the goal is reached, the metric is no longer useful because it implies a healthy process. Managers must be aware of this potential problem and use either supporting metrics or refurbished metrics to measure their processes.

Behavioral Extremes. The MATs derived from the demonstrations indicated that most of the behaviors

identified by the evaluation groups fell into Quadrants A and D. In other words, most of the behaviors were either clearly appropriate or clearly inappropriate. Some of this might have been caused by the directions presented by the research team. Although all behaviors were requested from the evaluation group, the research team emphasized the desire to include positive AND negative behaviors during the brainstorming session. This direction might have skewed them into the extreme positions. Also, the results might indicate that people have a stronger recall for positive and negative behavioral extremes. In other words, their short term memories have a capacity for remembering the extremes. Resolving this issue is beyond the scope of this research, however, future research should examine the dilemma.

The Assessment Process Continues. Based on the research team's analysis of the improvement actions, the assessment process is incomplete. Many of the improvement actions were anti-TQM in that they were police-type actions. Although TQM tries to avoid micromanagement, the improvement actions generated by the participants appeared to establish controls which would micromanage their behavior. The research team believes that the deficient behaviors should be assessed to determine which behaviors are the most damaging. After this is done, improvement actions can be generated and discussed under a TQM environment. The

evaluation team should identify those actions which might be micromanagement actions and discuss possible alternatives. Numerous checks and balances throughout the process can increase processing time and immobilize the workforce.

Metrics and Objectives. According to our literature review, metrics should support an organization's objectives. A measurement that does not motivate behaviors which will accomplish an objective is not an effective metric. In the thesis demonstrations, the evaluation groups assessed metric behaviors in terms of how well they supported the organizations objective. To prepare for this research, and during the research process, the research team discovered two basic concerns.

First, the B-2 metric focal point could not identify a specific objective to associate with the B-2 metric. They knew they wanted to decrease the number of days between the receipt of the ECP/CCP and the CCB action, but they could not identify why this was important. Upon further questioning, the B-2 focal point provided an implicit objective which was eventually tied to a HQ AFMC command objective. The basic objective provided was to "minimize the time it takes to incorporate an idea into a contract." Upon discussion, the B-2 group members indicated that the metric was developed after the number of contract changes began to dramatically increase. Thus, the group believed

that the objective of the metric was to reduce the overall number of contract changes in-house. In either case, it is clear that there was some confusion over the intended purpose of the measurement. After all, what good is it to measure the time it takes to approve a contract change proposal if the objective is to decrease the number of contract changes in house? Metrics should be born of objectives to ensure adequate motivation, measurement, and support for the organizational objectives.

Secondly, neither the B-2 nor the S&T demonstrations indicated a strong relationship between the metric and the organizational objective. The evaluation groups were asked to determine how well metric behaviors supported the organizational objective on a "1" (poorly) to "10" (very well) scale. Even when deficient behaviors were not considered, metric behaviors did not score above a "7." The research team believes that further research is in order to study the connection between the metrics and objectives more directly.

Identifying the Customer. The participants were required to assess the behaviors against the customer satisfaction attribute during the assessment process. Although the most critical point was to ensure that the participants were thinking of the same entity during the customer satisfaction scoring, defining the primary customer

was difficult. The research team was unsure of whether to test for the customer of the product or the customer of the output. The process owners (internal customers) may be concerned with process capacity or output, while the end users (external customers) would likely be concerned with the impact the process output has on the overall product. The primary customer was defined by the research team as the customer who benefits the most, and the most directly from the output of the process the metric measures. The research suggests behaviors that dissatisfied an output customer may have satisfied a product customer. Although all internal and external customers are important, the MAT is designed around a single customer. Therefore, the research team believes further research is required to determine the affect of different customers on the MAT and the process results. It may be that the MAT should be more of a multidimensional tool.

The GroupSystems Software. The GroupSystems Software V is useful for group discussions and decision making efforts. However, the research team believes that this, or any other software, is not required for a metric assessment. Many other tools and methods are available for brainstorming. The most critical factor is to ensure that ALL members are able to express their ideas freely, without criticism. The GroupSystem Software V system provided this opportunity and

generated instant feedback on all ideas and voting results. The evaluation group indicated that this software made their assessment experience more enjoyable.

## Recommendations for Future Research

The research team has six recommendations for continued research in the area of metrics.

1. A study of the MAT and assessment process use during metric development.
2. A direct study of the connection between metrics and objectives.
3. A comparison study of the MAT and assessment process using a variety of customers.
4. A study on whether the improvement actions have been implemented and whether they have been effective at controlling deficient behaviors.
5. A study to determine whether behavioral extremes are the only behaviors displayed within a process.
6. A comparison study of metrics used for process improvement (TQM) versus metrics based on the Theory of Constraints.

A Metric Development Technique. The MAT and assessment process could be used during metric development. Traditional metric development procedures include a step to assess a metrics effectiveness. The newly developed metric (or existing metric) can be tested using the MAT and the assessment process. From these techniques, personnel could discover whether the metric would motivate behaviors which are conducive to the organizational objective. They can

also determine the need for supporting metrics or other management actions which would improve the overall effectiveness of the metric. It would be an advantage to correct problems within the metric before it is implemented.

Metrics and Objectives. This study identified several concerns relating to the connection of metrics to objectives. The research team believes that people are not clear on the ultimate purpose of measurement within their processes. A direct study should be done to determine the strengths and weaknesses of the metric-objective relationship, and how this relationship can be enhanced. It would also be interesting to question employees to determine how many metrics existed that do not have clear and specific objectives associated with them. We believe that the opinions will vary significantly.

Customer Variety. The research team recommends that a study be done to determine the result different customer perspectives have on the MAT. Specifically, we believe a study is warranted to examine the difference between the customer of the process output, and the customer of the product the process supports. Behaviors, such as "working harder," might be perceived differently by these different customers which significantly impacts MAT behavior placement.

Improvement Actions. The demonstration of the MAT and the assessment process did not include a follow-on study to determine whether or not the improvement actions were implemented and how effective they were at improving the process measurement. This is a crucial component for determining the effectiveness of the MAT/assessment process. We believe that the improvement actions will be effective because they were derived directly from the process workers. However, a follow-on study should be done to determine if management has implemented any of the recommendations. If actions have been implemented, the metric could be evaluated with the MAT/assessment process once again to determine if the process measurement has been improved.

Behavioral Extremes. This research suggested that most of the behaviors occurring within the process were either inappropriate or appropriate. This may be indicative of the actual process behaviors. Researchers could study a process and record actual workforce behaviors. The actual behaviors could then be placed on the MAT using the assessment process. This would serve two purposes. First, the researchers could determine if behavioral extremes are realistic and indicative of the actual behaviors. Second, the researchers could assess the metrics for overall effectiveness.

TQM vs The Theory of Constraints. Many of the metrics developed for process improvements under TQM conflict with the principles of the Theory of Constraints (TOC). The TOC stipulates that improvements to nonconstraint processes might place further burdens on constraint resources and cause further declines in throughput. In other words, we might actually be doing our organizations more harm than good by improving **all** processes. It would be interesting to study whether or not a process is a constraint, then study the effects of process improvements on the TOC measurements of throughput, inventory, and operating expense.

## Summary

Although much literature exists concerning metric development, no literature exists concerning metric assessments. Every development process reviewed contained a "metric assessment" step, yet none offered a method for doing so. This research provides a foundation for metric assessment. It is clear much work remains in this area, but it is necessary that research continue. The United States Air Force is struggling to implement quality management. Much of its struggle is found in its measures. We cannot realize quality unless we know how to measure it. Viewing metrics as motivators, customer satisfiers, and as a vehicle for process improvement, will ensure that a metric is developed which will further the implementation of TQM. We

cannot measure everything and we can no longer afford to measure inappropriately. We must ensure that our measures are moving us in the proper direction, the direction towards customer satisfaction and process improvement.

## Appendix A: Definitions

Advanced Technology Transition Demonstration (ATTD): A laboratory project in a 6.3 Advanced Technology Development-funded program with the specific objective of meeting the user's defined needs through risk reducing "proof of principle" demonstrations conducted at the subsystem or higher level in an operationally realistic environment. (16:29)

Appropriate Behaviors: Actions which are directed towards an improvement to the process (5:1).

Behavior: Actions or responses taken by an organism in reaction to a given stimulus (38:100).

Continuous Improvement: Continuously monitoring processes to determine if they function as desired and if they can be improved (4:B-2). Undertaking improvement projects range from fixing things that fail to creating new processes, services, and products. It means solving a customer's immediate problems and it means preventing the same problem from happening again. (22:13).

Customer: Anyone who is impacted by an organization's processes and products. Internal and external (20:328).

Improvement Actions: Actions done in order to enhance the value or quality of a process. These are appropriate behaviors.

Measurement: The dimension, quantity or capacity determined by measuring (4:B-5).

Metric: A mathematical function which measures some aspect of the input and/or output of a process over time such that improvement actions to that process can be taken and the effects of those actions verified (5:1, 5).

Outcome: Something that follows as a result or consequence (38:814).

Process: A group of sequential, logically related tasks that use organizational resources to provide a product or a service to internal or external customers. The transformation of inputs into outputs (4:B-7).

Process Change: A transformation of a process.

Process Improvement: Includes any action done within the process to enhance the output.

Process Measurement: The accumulation and analysis of data used to assess the status of a process. Metrics are used to accumulate the data required.

Quality: The act of providing customers with products and services that consistently meet or exceed their needs and expectations (4:B-8).

Satisfaction: The fulfillment of a need or want (38:1026).

Sustaining Behaviors: Actions done in the short term to satisfy a customer or a measurement goal imposed by management. These actions do not change or improve a process (5:23).

Technology Transition Plan (TTP):
An agreement between the laboratory, the product division technology transition focal point, and the technology recipient that documents the specific tasks that must be successfully completed prior to technology acceptance. MAJCOM signatures are required if the MAJCOM ranked the ATTD 9.5 or higher. (16:29)

Temporary Actions: Actions which do not involve a process change towards improvement.

Total Quality Management: A management style which focuses on satisfying customer expectations by continually improving the way business is conducted.

## Appendix B: AFMC Command Goals and Objectives

**GOAL 1: SATISFY OUR CUSTOMER'S NEEDS...IN WAR AND PEACE**

Objective 1.1: Understand, through sustained interaction, our customers and their requirements, and provide options, including those available through other services which are the basis for customer decisions and satisfaction.

Objective 1.2: Ensure a robust AFMC warfighting posture, including transition from peace to war.

Objective 1.3: Be our customers' supplier of choice by: meeting cost, schedule and performance baselines; enhancing customer support and lowering life cycle costs.

Objective 1.4: Meet anticipated customer needs by planning for and securing continuing support of capital investments in AFMC infrastructure.

**GOAL 2: ENABLE OUR PEOPLE TO EXCEL**

Objective 2.1: Create, implement and communicate a career development program for all military and civilian personnel in the command.

Objective 2.2: Invest in our people by providing necessary education and training.

Objective 2.3: Move decisions to the lowest level, expand individual responsibility and authority, and seek and provide feedback.

Objective 2.4: Champion and implement personnel management changes that enhance productivity and job satisfaction.

Objective 2.5: Optimize the workforce mix to conduct the AFMC mission.

**GOAL 3: SUSTAIN TECHNOLOGICAL SUPERIORITY**

Objective 3.1: Continuously improve quality and relevance of Air Force laboratory science and technology programs.

Objective 3.2: Transition technology rapidly to applications to include organic infrastructure.

Objective 3.3: Leverage the science and technology investment of other defense and government labs, allies , academia and industry.

**GOAL 4: ENHANCE THE EXCELLENCE OF OUR BUSINESS PRACTICES**

Objective 4.1: Enhance the competitiveness of our operations by improving throughput, and decreasing inventory and operating expense in everything we do.

Objective 4.2: Expand and mature Integrated Weapon System Management (IWSM) by continuously improving teamwork, policies and processes.

Objective 4.3: Champion environmental responsibility in design, test, support and industrial processes.

Objective 4.4: Increase our business with high quality suppliers to control and improve delivered products and services at all values.

Objective 4.5: Pursue joint solutions, interservicing and interoperability in our business practices.

**GOAL 5: OPERATE QUALITY INSTALLATIONS**

Objective 5.1: Enhance the quality of life of our people through continuous improvement in facilities, infrastructure, services and work environment to satisfy their needs and priorities.

Objective 5.2: Be a good neighbor by enhancing community relationships.

Objective 5.3: Demonstrate environmental leadership through proper planning and execution of restoration, compliance and hazardous waste disposal programs.

## Appendix C: The Proposed Metric Assessment Process

METRIC EXAMPLE: Number of units delivered vs. the number of units required to be delivered on a monthly basis.

**STEP ONE: IDENTIFICATION OF BEHAVIORS**

Question: What behaviors might occur as a result of this metric?

Examples: Increase amount of overtime to meet required deliveries
Increase emphasis on project planning
People will be more concerned with quantity than quality.

**STEP TWO: DISCUSSIONS/REFINE BEHAVIORS**

**STEP THREE: IDENTIFY CUSTOMER**

Question: Identify all customers (both internal and external) who might have an interest in the output of the process being measured.

Examples: Supervisor
End Users

* Vote on Primary

**STEP FOUR: CRITERIA EVALUATION**

Question: If you were the primary customer, would you be satisfied with this behavior? (Y or N)

Question: Is the behavior directed towards an improvement to the process? (Y or N)

**STEP FIVE: EXPLANATION OF METRIC ASSESSMENT TOOL**

**STEP SIX: EVALUATION OF ORGANIZATIONAL OBJECTIVE**

Question: How well do the composite behaviors support the organizational objective? (Rating: poorly to very well)

Question: How well do the quadrant A behaviors support the organizational objective? (Rating: poorly to very well)

**STEP SEVEN: BRAINSTORM IMPROVEMENT ACTIONS**

Question: What are your recommendations for improving the process measurement?

Examples: Get approval before overtime can be used.
Support metric needed to account for quality.

**Step EIGHT: DISCUSS/REFINE IMPROVEMENTS**

**FINAL: QUESTIONNAIRE**

## Appendix D: Software Reference Guide

**IDENTIFICATION OF BEHAVIORS:**

Step 1: Press the **insert** key to enter the first behavior.
Step 2. **Enter behavior** (one liner).
Step 3: (**Enter**).
Step 4: **Input Comments**. (if desired)
Step 5: **F3** to send behavior to master list.
Step 6: **Return to Step 1**.

**VOTING ON PRIMARY CUSTOMER:**

**Rearrange List** of customers by listing your first choice for primary customer first, your second choice second,...etc., by:

Step 1: **Place** the **cursor** on the customer you want to move.
Step 2: **hit** the **space bar**.
Step 3: **move** the **cursor** to where you want that customer to be placed, and
Step 4: **hit** the **space bar** again.
Step 5: **Repeat** steps **until** your rearrangement is **complete**.
Step 6: Hit **F3** to signal your completion and register your votes.
Step 7: Cast Vote? Y/N Hit **Y** if you are ready to cast your vote.

**EVALUATING BEHAVIORS:**

**CRITERIA I**: The first behavior and the first criteria (customer satisfaction) will be at the top of the screen.

Step 1: Hit (**Enter**) to bring up **rating scale**. (1 to 10)
Step 2: In accordance with the criteria, rate the behavior by **moving the cursor** (using arrow keys) **over the number** you wish to select and pressing (**Enter**) casting your vote.
Step 3: Go to Step 2 to **rate the next behavior** in accordance with the first criteria.

| | |
|---|---|
| **SCALE**: | 10 = YOU WOULD BE EXTREMELY SATISFIED WITH THIS BEHAVIOR. |
| | 1 = YOU WOULD BE EXTREMELY DISSATISFIED OR NOT AFFECTED BY THIS BEHAVIOR |

Step 4: Hit **F3** to cast your votes and review the results.

The second criteria will automatically appear on your screen along with the first behavior.

**CRITERIA II**: The second criteria will appear (improving action) along with the first behavior.

Step 5: Cast your **VOTE** for criteria two. (ref. Step 2)

The next behavior will automatically appear on your screen. **Begin again with Step 2** until all votes are complete.

Step 6: Your rating summary will appear on the screen. Hit **Escape** to continue.
Step 7: Hit **F3** to cast your votes and review the results.
Step 8: Review rating summary? Y/N Hit **N** to continue.
Step 9: Cast Vote and exist? Y/N Hit **Y** if you are ready to cast your votes.

| | |
|---|---|
| **SCALE**: | 10 = THE BEHAVIOR DEFINITELY CONTRIBUTES TO PROCESS IMPROVEMENT |
| | 1 = THE BEHAVIOR DEFINITELY DOES NOT CONTRIBUTE TO PROCESS IMPROVEMENT |

**EVALUATING THE ORGANIZATIONAL OBJECTIVE:**

Step 1: **Rate** the degree to which the composite behaviors support the organizational objective on a scale of 1 to 10. Move the cursor to the number you wish to select.
Step 2: Hit (**Enter**).

Step 3: **Rate** the degree to which the quadrant A behaviors support the organizational objective on a scale of 1 to 10. Move the cursor to the number you wish to select.
Step 4: (**Enter**).
Step 5: Hit **F3** to signal completion of votes.
Step 6: Cast vote and exist? Y/N Hit **Y** to cast your vote.

| |
|---|
| **SCALE**: 1 = VERY WELL<br>10 = POORLY |

**IDENTIFICATION OF MEASUREMENT IMPROVEMENT ACTIONS**:

**PHASE I**: Actions for **Individual Behaviors**.

Step 1: Look at each individual behavior in quadrants B, C, and D. Make recommendations which will control or eliminate the behavior where possible. Follow Steps 2 - 7.

**PHASE II**: Actions for **Metric Structure**

Step 1: Look at the behaviors generated by the metric in general, make any recommendations for changes to the structure of the metric or add any supporting metrics you wish. Follow Steps 2 - 7.

**PHASE III**: **Other Actions**.

Step 1: Make any other recommendations you might have to improving the measurement of the process. Follow Steps 2 - 7.

Step 2: Hit **Insert** to begin adding recommendations.
Step 3: Make the **Recommendation**. (one liners)
Step 4: (**Enter**).
Step 5: **Provide explanation** or clarification **in comment section**. **(Be as specific as you can)**
Step 6: Hit **F3**.
Step 7: Add **next recommendation**. Return to Step 2.

## Appendix E: The Metric Assessment Process

METRIC EXAMPLE: Number of units delivered vs. the number of units required to be delivered on a monthly basis.

**STEP ONE: IDENTIFICATION OF BEHAVIORS**

Question: What behaviors might occur as a result of this metric?

Examples: Increase amount of overtime to meet required deliveries
Increase emphasis on project planning
People will be more concerned with quantity than quality.

**STEP TWO: DISCUSSIONS/REFINE BEHAVIORS**

**STEP THREE: PRIMARY CUSTOMER VOTING**

**STEP FOUR: CRITERIA EVALUATION**

Question: As the primary customer, characterize your satisfaction with each behavior. (Rating: 1 to 10)

Question: Determine the degree to which each behavior contributes to process improvement. (Rating: 1 to 10).

**STEP FIVE: EXPLANATION OF METRIC ASSESSMENT TOOL**

**STEP SIX: EVALUATION OF ORGANIZATIONAL OBJECTIVE**

Question: How well do the composite behaviors support the organizational objective? (Rating: 1 to 10)

Question: How well do the quadrant A behaviors support the organizational objective? (Rating: 1 to 10)

**STEP SEVEN: IDENTIFICATION OF MEASUREMENT IMPROVEMENT ACTIONS**

**PHASE I: Individual Behaviors**

Question: Focus on the behaviors that are undesirable. What management actions might you recommend to control or redirect these behaviors.

**PHASE II: The Metric Structure**

Question: Consider the metric as it is currently written. What structural changes or supporting metrics would you recommend to improve the process measurement?

**PHASE III: Other**

Question: Are there any other recommendations you might have to improve the process measurement?

**STEP EIGHT: IMPROVEMENT ACTION DISCUSSIONS/CLARIFICATIONS**

## Appendix F: The Metric Assessment Fact Sheet - S&T

ORGANIZATIONAL OBJECTIVE 3.2: Transition technology rapidly to applications, to include organic infrastructure.

**TEAMWORK METRIC**: The percent of ATTD's budgeted for the upcoming fiscal year with signed TTP's

PROCESS: Preparation and approval (signatures from ASC/EN, Wright-Laboratory, and the specified SPO (also the MAJCOM if necessary)) of TTP's for each ATTD.

OUTPUT: A signed TTP

PRIMARY CUSTOMER: _____

PARTICIPANT NOTES:

## Appendix G: The Metric Assessment Fact Sheet - B-2

ORGANIZATIONAL OBJECTIVE: Minimize the time it takes to incorporate an idea (for product enhancement or deficiency correction) into a contract.

**RESPONSIVENESS METRIC**: The average number of days a prepared ECP takes to be approved by the CCB

PROCESS: Preparation of ECPs and the approval of the engineering change by the CCB

OUTPUT: An approved ECP

PRIMARY CUSTOMER: _____

PARTICIPANT NOTES:

## Appendix H: Outcomes and "It Depends" Syndrome

Metric Behaviors/Outcomes Example

| BEHAVIORS | OUTCOMES (NON-BEHAVIORS) |
|---|---|
| - Increase amount of overtime to meet required deliveries<br><br>- Increase emphasis on project planning<br><br>- People will be more concerned with quantity than quality<br><br>- An individual may fill several orders at one time<br><br>- Change delivery schedule | - All units will be delivered on time<br><br>- Metric brought process sluggishness to management's attention<br><br>- 2-Letter management attention increases time to complete process<br><br>- Metric does not measure continued improvement to the process |

"It Depends" Example

Metric Behavior: An individual may fill several orders at one time

| POSITIVE INTERPRETATION | NEGATIVE INTERPRETATION |
|---|---|
| - Filling several orders at once streamlines operations and speeds up delivery to the customer | - Filling several orders at one time may put quantity over quality and ultimately result in reworking orders, customer complaints etc. |

# Appendix I: S&T Metric Assessment Tool

| | | |
|---|---|---|
| | **A** | **B** |
| **Customer Satisfaction** | 2. Personnel motivated to complete TTP's in most cases<br><br>3. Increased desire by TTP author to understand process<br><br>4. metric brought process sluggishness to management attention<br><br>5. Two-letter management pays more attention to quality in process results<br><br>9. Forced TTP author to interface/delegate duties- work smarter<br><br>11. Made process players more involved (i.e. boss is watching)<br><br>14. Personnel encouraged to interact with customer | 8. Forced TTP author to handcarry document at times - work harder |
| **No Customer Satisfaction** | **C** | **D** |
| | 7. TTP signature level process is faster | 1. Personnel (all levels) emphasize quantity over quality<br><br>6. 2-letter management attention increases time to complete process<br><br>10. Personnel emphasize quantity over process<br><br>12. TTP author was often unaware of WL/XPT's schedule of events<br><br>13. Metric measured "one time" get well program<br><br>15. Forced TTP schedules create encouragement to "get it done"<br><br>16. TTP author at directorate level felt lack of authority over process |
| | **Improving Actions** | **Sustaining Actions** |

## Appendix J: B-2 SPO Metric Assessment Tool

| | | |
|---|---|---|
| **Customer Satisfaction** | **A**<br>6. Would try to do more up front coordination to facilitate process<br>7. Could cause you to seek proposal review comments before suspense due (+)<br>11. May increase team motivation/cooperation/creativity<br>14. Would lead to creative ways to define requirements and deal with the contractor<br>16. Could cause one to ask contractor for proposal change pages (+)<br>20. Encourages modification of work habits/increases efficiency | **B**<br>2. Encourages putting the measured proposal as the top priority (+) |
| **No Customer Satisfaction** | **C** | **D**<br>1. Effort expended to get around the process instead of using it correctly<br>3. Could cause one to neglect other work<br>4. People concerned with due dates rather than quality technical proposal<br>5. Managers may use the metric as a personnel evaluation tool (-)<br>8. Might cause one to schedule CCB date prematurely<br>9. Try to do "CYA" activities, like blame the contractor<br>10. May cause "unhealthy competition" between the teams<br>12. May induce conflict between process owners and process users<br>13. Try to have proposal designated as number one priority in SPO<br>15. Lead to inappropriate actions in dealing with the contractor<br>17. May cause people to "fudge" numbers to make metric look good<br>18. Could cause extended work hours (work harder)<br>19. Could cause unnecessary TDY to correct deficiencies with contractor |
| | **Improving Actions** | **Sustaining Actions** |

## Appendix K: The Metric Assessment Tool Explanation

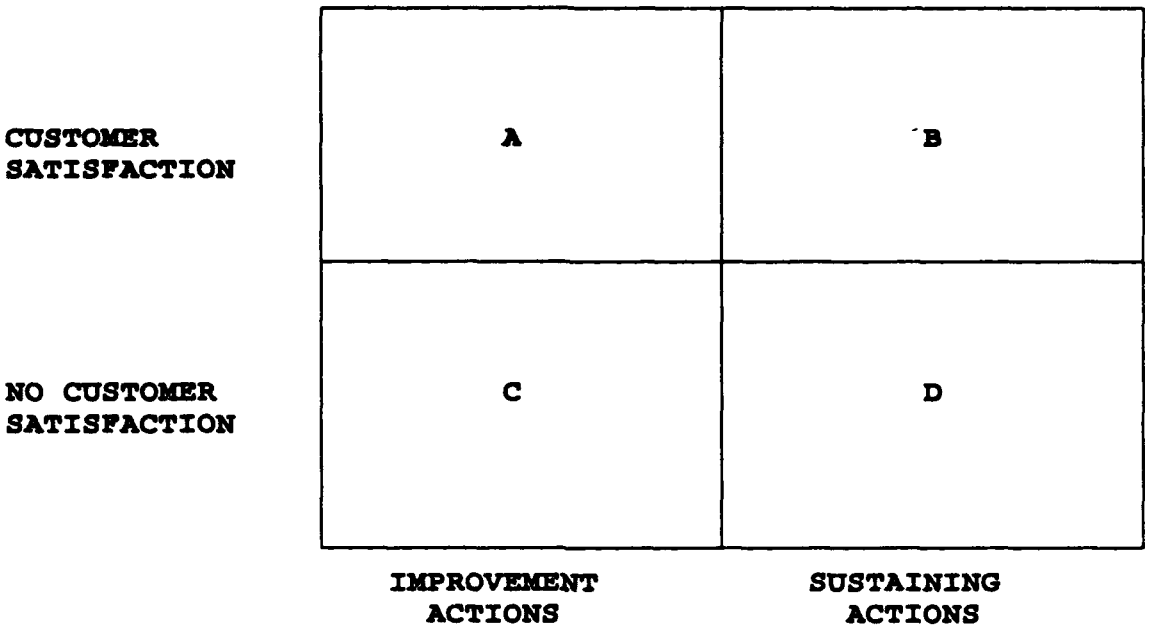



**FIGURE 3.1 THE METRIC ASSESSMENT TOOL**

Quadrant A. Behaviors that fall within Quadrant A reflect a motivation towards process improvements which will result in customer satisfaction. As such, these behaviors reflect the ultimate goal of a good metric (i.e., improving the process and customer satisfaction). Quadrant A also provides insight into whether or not the behaviors generated by the metric will support an organization's objective. For example, if the organization's objective was to decrease the number of defective parts produced and a behavior identified in Quadrant A was to emphasize inspection at critical inspection points, one may contend that the metric, and thus the process measurement, are promoting the objective.

Quadrant B. Behaviors that fall within Quadrant B are behaviors which are not directed towards process improvements. They are generally temporary actions done as an immediate response to ensure customer satisfaction or to meet some particular measurement goal. These behaviors may ultimately distort the process measurement. For example, if a measured process indicates that 50 contracts were awarded within 100 days. The customer would be satisfied because he has received his award. However, a closer look into the

process reveals that people stayed until midnight in order to award 5 of the 50 contracts. The process measurement is distorted because it no longer reflects the actual results of the process. The process only generated 45 contracts but the extra hours made it appear as though the process was generating more.

Quadrant C. Behaviors that fall within Quadrant C are behaviors which are directed towards process improvement which do not result in customer satisfaction. For example, if a metric motivates a process change for employees to account for their efforts on a per-project level, the process change may not influence customer satisfaction. As a result, the behaviors would not necessarily support the delivery of quality products and services and may not be considered a valued improvement.

Quadrant D. Behaviors falling in Quadrant D are similar to the behaviors in Quadrant B in that they may ultimately distort the process measurement. In addition, behaviors falling in Quadrant D will not result in customer satisfaction. For example, a customer wants delivery in 30 days. The metric measures the actual delivery days to the scheduled delivery days as a ratio. One behavior which might be generated by this metric would be a modification of the scheduled delivery date to the customer. Although the customer may not have a choice but to agree to this new date, he will probably be dissatisfied. The behavior did not result in customer satisfaction, nor did it motivate process improvement.

## Appendix L: Post Process Questionnaire

Please respond to each of the questions below. Any comments which will improve the overall effectiveness of the Metric Assessment Tool (MAT) and/or its process are encouraged and welcomed. All answers will be considered confidential.

**The Metric Assessment Tool:**

1. The metric assessment tool (was/was not) easy to understand. WHY?

2 In my opinion, the MAT (did/did not) enhance my understanding of the relationship between the metric and the process? WHY?

3. I personally found the MAT to be useful for metric assessments? (Yes or No) WHY?

4. I (would/would not) recommend the MAT to others wishing to evaluate metrics. WHY?

5. Did you have any difficulty answering the questions regarding customer satisfaction, process improvement, and/or objective support? (Yes or No) WHY?

6. How many of the behaviors identified have you personally witnessed within the process? Circle your best answer.

All Many Some Few None

**The Metric Assessment Process:**

7. The metric assessment process (was/was not) enhanced by the use of the computer software system at Armstrong Laboratory. WHY?

8. Do you feel that the list of improvements you generated will improve the current approach to measuring the process? (Y or N) WHY?

**The Metric Assessment Tool and Process:**

9. Please provide any comments concerning the POSITIVE aspects of the MAT and/or the assessment process.

10. Please provide any comments concerning NEGATIVE aspects of the MAT and/or the assessment process.

11. Please provide any suggestions for improving the MAT and/or the assessment process.

12. Other comments.

## Appendix M: S&T Questionnaire Summary

The following summary represents the actual responses as received from each evaluation group member. No attempt has been made to alter the grammar or syntax of any response. The questions below are followed by the comments of each evaluation group member.

**The Metric Assessment Tool:**

**1. The metric assessment tool (was/was not) easy to understand.WHY?**

WAS. After a little thought - it was easy. Perhaps you can slow down a little - instead of reading use viewgraphs to restate what you want.

WAS. Provided an easy way to categorize inputs.

WAS. Never quite got a handle on the math that turned the scores into quadrant locations but I am sure it made sense.

**2. In my opinion, the MAT (did/did not) enhance my understanding of the relationship between the metric and the process? WHY?**

The categorization process was helpful.

It displayed how inputs might fall into certain categories.

DID NOT. In the case studied, I think the group focused on the process shortfalls and not on the metric.

**3. I personally found the MAT to be useful for metric assessments? (Yes or No) WHY?**

YES. Be careful though - I felt that we may have missed some potential pitfalls - i.e. options for other fixes to the metric.

YES. It showed how different inputs/responses would fall into different categories.

YES. Allowed organization of thoughts and ideas to be readily organized for discussion.

**4. I (would/would not) recommend the MAT to others wishing to evaluate metrics. WHY?**

WOULD. It causes you to "think - it - through" - group discussions were good.

WOULD. Evaluation tool - good approach

WOULD. See 3. (third comment)

**5. Did you have any difficulty answering the questions regarding customer satisfaction, process improvement, and/or objective support? (Yes or No) WHY?**

NO. But it did help when you repeated what you wanted several times.

YES. Not entirely familiar with topics and processes involved. Therefore, don't have a fine understanding of the issues.

YES. It was often difficult to establish relationship to customer or process.

**6. How many of the behaviors identified have you personally witnessed within the process? Circle your best answer.**

All Many Some Few None

ALL.

MANY.

MANY.

**<u>The Metric Assessment Process</u>:**

**7. The metric assessment process (was/was not) enhanced by the use of the computer software system at Armstrong Laboratory. WHY?**

WAS. This was especially useful/fun to use - good aide.

WAS. The software made the evaluation process easier to go through, follow and to provide inputs to.

WAS. Certainly, especially had our group been larger. In larger groups, personalities take over and not everyone has an opportunity to participate.

**8. Do you feel that the list of improvements you generated will improve the current approach to measuring the process? (Y or N) WHY?**

YES. Collectively I feel that something good will come of all of our work.

YES. Hope so, but I don't see HQ taking guidance and inputs very well from the working level.

YES. It will bring to the attention to management some of the problems with process management wants to improve, but sometimes can't see the forest through the trees.

**The Metric Assessment Tool and Process:**

**9. Please provide any comments concerning the POSITIVE aspects of the MAT and/or the assessment process.**

Group dynamics was good - got to see the "other side" of the question. The rescoring after discussion was extremely good -will help your quality.

Painless to follow.

Brought out constructive criticism of the process that may not otherwise be brought out.

**10. Please provide any comments concerning NEGATIVE aspects of the MAT and/or the assessment process.**

None really - the "fixes" to the system - perhaps we could have discussed these together.

Difficult to follow at times.

Sometimes hard to distinguish between positive and negative impact of behaviors (i.e., no middle of the road).

**11. Please provide any suggestions for improving the MAT and/or the assessment process.**

See # 10 above (first comment)

NONE

See 10 (third comment). Things aren't always black and white.

**12. Other comments.**

NONE

NONE

Nice facility. Will advertise it to others.

## Appendix N: B-2 SPO Questionnaire Summary

The following summary represents the actual responses as received from each evaluation group member. No attempt has been made to alter the grammar or syntax of any response. The questions below are followed by the comments of each evaluation group member.

**The Metric Assessment Tool:**

**1. The metric assessment tool (was/was not) easy to understand. WHY?**

WAS. Fairly easy to understand. Actually easier once we went through the process.

WAS. Well defined and explained. Courteous assistance and guidance played key role.

WAS. Very user friendly.

WAS. Everything was prepared and organized.

**2. In my opinion, the MAT (did/did not) enhance my understanding of the relationship between the metric and the process? WHY?**

Provided an opportunity to think about the relationship.

DID NOT. Already had a good understanding.

Forces a better understanding of what drives the results of the metric.

The distinction between improving actions and sustaining actions was explained.

**3. I personally found the MAT to be useful for metric assessments? (Yes or No) WHY?**

YES. Provided a forum for metric assessment.

YES. Gives you feedback on what other people think.

YES. Brings others' ideas into play other than your own.

YES. It reveals what is/is not satisfying to the customer.

**4. I (would/would not) recommend the MAT to others wishing to evaluate metrics. WHY?**

WOULD. Provides the discipline and procedure to effectively evaluate a metric.

WOULD. Excellent tool to determine if metric is meaningful to anyone but you and will achieve goals.

WOULD. Allows others to evaluate the usefulness of their metric and opens thought processes for improving.

WOULD. It's an easy to understand tool.

**5. Did you have any difficulty answering the questions regarding customer satisfaction, process improvement, and/or objective support? (Yes or No) WHY?**

NO.

NO. Clearly explained.

NO.

NO for customer satisfaction/process improvement.

YES for objective support

**6. How many of the behaviors identified have you personally witnessed with the process? Circle your best answer.**

All Many Some Few None

MANY.

ALL.

MANY.

MANY.

**The Metric Assessment Process:**

**7. The metric assessment process (was/was not) enhanced by the use of the computer software system at Armstrong Laboratory. WHY?**

WAS. Easier to record thoughts. Aided in concentration.

WAS. Facilitated and equalized input.

WAS. Forces in-depth meaning into process other than just meeting a schedule.

WAS. Easy to use. Explained.

**8. Do you feel that the list of improvements you generated will improve the current approach to measuring the process? (Y or N) WHY?**

YES. If adopted.

DON'T KNOW. Depends on management buy-in.

YES. Cause re-evaluation of our metric for improvement areas.

YES. If incorporated.

**The Metric Assessment Tool and Process:**

**9. Please provide any comments concerning the POSITIVE aspects of the MAT and/or the assessment process.**

Environment and means of data gathering encourage development of ideas anonymously.

NONE.

Instant feedback. Saves time over traditional brainstorming methods.

The 4 quadrants make it easy to see where the factors fall (which ones are desirable)

**10. Please provide any comments concerning NEGATIVE aspects of the MAT and/or the assessment process.**

Could use a more complete overview at the beginning.

Any assessments are subjective and must be evaluated with that perspective in mind.

NONE.

Only had one computer glitch and slight delay. No problem though.

**11. Please provide any suggestions for improving the MAT and/or the assessment process.**

NONE.

NONE.

NONE.

Good job.

**12. Other comments.**

None

Nice job by the folks giving the assessment.

NONE.

Well organized staff. Explained things step-by-step. Courteous. Nice. Free coffee and food. Fun time. Thanks.

## BIBLIOGRAPHY


1. Aeronautical Systems Division, Air Force Materiel Command. The ASD Metrics Handbook. ASD Pamphlet 700-8. Wright-Patterson AFB OH: ASD, 30 April 1992.

2. Aeronautical Systems Division. Improving the Winning Edge. Wright-Patterson AFB OH: HQ ASD, 1991.

3. Aeronautical Systems Division: TQ Office and the Cumberland Group. Measurement: The Handle on your Business! Wright-Patterson AFB OH, March 15, 1991.

4. Air Force Logistics Command. Quality Participation for Employees. Wright-Patterson AFB OH: HQ AFLC, (undated).

5. Air Force Materiel Command. The Metrics Handbook - Draft. AFMC Pamphlet 74-9. Wright-Patterson AFB OH: HQ AFMC/XPD, 26 March 1993.

6. Armstrong Laboratory. Meeting More and Enjoying it Less? Information Pamphlet. Armstrong Laboratory, Logistics Research Division, Wright-Patterson AFB OH: AFMC/AL/HRGA, (undated).

7. Bemowski, Karen. "The Benchmarking Bandwagon," Quality Progress: 19-24 (January 1991).

8. Bond, David F. "Military Budget Drawdown Spurs Movement Toward Increased Efficiency," Aviation Week & Space Technology: 58-59 (December 1991).

9. Cream, Bertram W. and Lt Col Charles J. Botello. "Armstrong Laboratory: Mission Briefing." Report to AFMC, OL-AL/HRG, Wright-Patterson AFB OH: January 1993.

10. Deming, W. Edwards. "Improvement of Quality and Productivity through Action by Management," National Productivity Review: 12-22 (Winter 1981-82).

11. Engen, D. Travis. "Total Quality Management," Air Force Journal of Logistics: 2-3 (Winter 1990).

12. Finison, Lorenz J. "What Are Good Health Care Measurements?" Quality Progress: 41-42 (April 1992).

13. Gitlow, Howard S., and Shelly J. Gitlow. The Deming Guide to Quality and Competitive Position. Englewood Cliffs NJ: Prentice-Hall Inc., 1987.

14. -----. Handbook of Small Group Research. New York: The Free Press, 1976.

15. Harrington, H.J. The Improvement Process. Milwaukee WI:McGraw-Hill, 1987.

16. HQ AFMC/ST. "Air Force Science & Technology Metrics." Report to Headquarters, HQ AFMC, Wright-Patterson AFB OH, 1 April 1993.

17. HQ AFMC/YPX. "Toward New Horizons." Pamphlet, HQ AFMC, Wright-Patterson AFB OH, (undated).

18. Jessup, Leonard M. and David A. Tansik. "Decision Making in an Automated Environment: The Effects of Anonymity and Proximity with a Group Decision Support System," National Productivity Review, 8: 145-156 (Spring 1989).

19. Jewell, Linda N. and H. Joseph Reitz. Group Effectiveness in Organizations. Glenview IL: Scott, Forseman and Co., 1981.

20. Juran J.M. Juran on Planning for Quality. New York: The Free Press, 1988.

21. Kaplan, Robert S., and David P. Norton. "The Balanced Scorecard-Measures That Drive Performance," Harvard Business Review: 71-79 (January/February 1992).

22. Kinlaw, Dennis C. Continuous Improvement and Measurements for Total Quality. Homewood IL: Pfeiffer & Company, 1992.

23. Lynch, Richard L., and Kelvin F. Cross. Measure Up! Yardsticks for Continuous Improvement. Cambridge: Blackwell Publishers, 1992.

24. McPeak, Merrill A. General, Chief of Staff of the Air Force, The Pentagon, Washington DC. Official Correspondence. 24 December 1992.

25. Michaelson, Larry K., Warren E. Watson, and Robert H. Black. "A Realistic Test of Individual Versus Group Consensus Decision Making," Journal of Applied Psychology 74: 834-839 (1989).

26. Miller, Lawrence M., and Robert J. Hayes. An Evaluation of Schedule Metrics Used Within Aeronautical Systems Center. MS Thesis, AFIT/GSM/LAA/92S-12. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1992. (AD-A260113)

27. Moore, Carl M. Group Techniques for Idea Building. Beverly Hills CA: Sage Publications Inc., 1987.

28. Munoz, Jairo, and Chase Nielsen. "SPC: What Data Should I Collect? What Charts Should I Use?" Quality Progress: 50-52 (January 1991).

29. Office of Management and Budget (OMB). How to Develop Quality Measures That are Useful in Day-To-Day Management. Washington: Government Printing Office, January 1989.

30. Peasant, Janet L. "Group Support Technology," S&T ORT Briefing, AL/HRGA, Wright-Patterson AFB OH: 3 Feb 93.

31. Robinson, James D. "An Open Letter: TQM on Campus," Harvard Business Review : 94 (Nov/Dec 1991).

32. Schaffer, Robert H., and Harvey A. Thomson. "Successful Change Programs Begin with Results," Harvard Business Review 80: 80-89 (January/February 1992).

33. Showalter, Michael J., and Judith A. Mulholland. "Continuous Improvement Strategies for Service Organizations,." Business Horizons : 82-87 (July - August 1992).

34. Swiss, James E. "Adapting Total Quality Management (TQM) to Government,"Public Administration Review 52: 356-360 (July/August 1992).

35. United States Office of Personnel Management (USOPM). Federal Total Quality Management Handbook. Series 3; Washington: Government Printing Office, May 1989.

36. Walton, Mary. The Deming Management Method. New York: The Putnam Publishing Group, 1986.

37. Walton, Mary. Deming Management at Work. New York: The Putnam Publishing Group, 1990.

38. Merriam-Webster. Websters New Collegiate Dictionary. Springfield MA: G & C Merriam Company, 1973.

## Vita

Cynthia A. Campbell was born on 24 March 1960 in Cleveland Ohio. She graduated from The Ohio State University in March 1983 where she received a Bachelor of Science in Business Administration Degree. While at Ohio State, Ms. Campbell majored in Human Resource Management. After graduating, Ms. Campbell was hired by the Air Force under the Pacer Produce I employment action. After extensive training by Wright Patterson AFB, Ohio, Ms. Campbell was transferred to Warner Robins Air Logistics Center where she was employed as a Contract Price Analyst. While at Warner Robins, she established and managed an independent defective pricing organization. Her abilities were recognized with a transfer to the Headquarters Air Force Logistics Command where she was the primary action officer on several pricing policy issues. She entered the School of Systems and Logistics at the Air Force Institute of Technology in May 1992 to pursue her masters degree in contracting. Upon graduation, she will be assigned to the Aeronautical Systems Center to pursue her career in contracting.

Permanent Address: 1135 Katie Circle
Beavercreek, Ohio 45434

## Vita

Captain Gregory M. Gutterman was born on 2 February 1966 in St. Paul, Minnesota. He graduated from Park High School in Cottage Grove, Minnesota in 1984 and attended the United States Air Force Academy, graduating with a Bachelor of Science in Engineering Mechanics (specialty: Structures) in June 1989. Upon graduation, he received a regular commission in the USAF and served his first tour of duty at Grand Forks AFB, North Dakota. As a Technical Engineer for the 321st Missile Wing, he provided hands on guidance to maintainers, advice to the Deputy Commander for Maintenance, and initiated several modifications to the Minuteman III weapon system. He was then assigned to the Air Force Institute of Technology as a graduate student of Systems Management in May 1992.

Permanent Address: 616 Hickory Street
Dayton, Ohio, 45410

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | September 1993 | Master's Thesis |

**4. TITLE AND SUBTITLE**
THE DEVELOPMENT AND DEMONSTRATION OF THE METRIC ASSESSMENT TOOL

**5. FUNDING NUMBERS**

**6. AUTHOR(S)**
Cynthia A. Campbell, GS-13, USAF
Gregory M. Gutterman, Captain USAF

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Air Force Institute of Technology,
WPAFB OH 45433-6583

**8. PERFORMING ORGANIZATION REPORT NUMBER**
AFIT/GCM/LAS/93S-3

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
HQ AFMC/XPD
WPAFB OH 45433-6583

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*
This study reflects the development and demonstration of the Metric Assessment Tool. The purpose of the tool was to provide individuals the means to assess metrics and make improvements to the process measurement. The tool was developed using two critical attributes: customer satisfaction and process improvement. Once the tool was developed, a metric assessment process was designed to demonstrate the tool. Two metrics were selected from the Aeronautical Systems Center (ASC) and individuals with a working knowledge of the metric and process were selected for the demonstrations. Using a Group Support System at Armstrong Laboratory, Wright Patterson AFB, the group was asked to identify behaviors which might be motivated from the metric. Once the behaviors were identified, the group evaluated the behaviors against the critical attributes. From this assessment, behaviors were placed on the Metric Assessment Tool. This tool clearly identified deficient behaviors and how they might distort the process measurement. From this information, the group was asked to generate improvement actions which would serve to eliminate or control deficient behaviors. With the elimination or control of deficient behaviors, the process measurement is improved and the organizational objective is better served.

| 14. SUBJECT TERMS | 15. NUMBER OF PAGES |
|---|---|
| Work Measurement, Performance, Motivation, Measuring Instruments Metrics, Performance Measures, Total Quality Management | 163 |
| | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UL |

AFIT Control Number AFIT/GCM/LAS/93S-3

# AFIT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFIT thesis research. Please return completed questionnaires to: DEPARTMENT OF THE AIR FORCE, AIR FORCE INSTITUTE OF TECHNOLOGY/LAC, 2950 P STREET, WRIGHT PATTERSON AFB OH 45433-7765

1. Did this research contribute to a current research project?

   a. Yes  b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFIT had not researched it?

   a. Yes  b. No

3. The benefits of AFIT research can often be expressed by the equivalent value that your agency received by virtue of AFIT performing the research. Please estimate what this research would have cost in terms of manpower and/or dollars if it had been accomplished under contract or if it had been done in-house.

   Man Years _____  $ _____

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3, above) what is your estimate of its significance?

   a. Highly Significant  b. Significant  c. Slightly Significant  d. Of No Significance

5. Comments

_____ _____

Name and Grade  Organization

_____ _____

Position or Title  Address

DEPARTMENT OF THE AIR FORCE
AFIT/LAC Bldg 641
2950 P St
45433-7765

OFFICIAL BUSINESS

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

# BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 1006 DAYTON OH

POSTAGE WILL BE PAID BY U.S. ADDRESSEE

Wright-Patterson Air Force Base

AFIT/LAC Bldg 641
2950 P St
Wright-Patterson AFB OH 45433-9905