AFIT/GSM/LAS/93S-6

AN EXPLORATORY SURVEY OF METHODS USED TO DEVELOP MEASURES OF PERFORMANCE

THESIS

Kenneth L. Hamner, Captain, USAF

Charles A. La Fleur, Captain, USAF

AFIT/GSM/LAS/93S-6

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/GSM/LAS/93S-6

AN EXPLORATORY SURVEY OF METHODS USED TO DEVELOP MEASURES OF PERFORMANCE

THESIS

Presented to the Faculty of the

Graduate 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 Systems Management

Kenneth L. Hamner, B.S.

Captain, USAF

Charles A. La Fleur, B.S.

Captain, USAF

September 1993

Acknowledgments

Ken and Chuck thank those persons who contributed to the success of this research effort. We thank our advisors, Majors Scott Graham and Kevin Grant, for their willingness to go the extra mile on our behalf. They kept up with several research efforts simultaneously, and still found the time to perform their normal faculty responsibilities.

Thanks also to the members of the Faculty Qualification and Recruitment Committee, whose willingness to assist this research effort even after a full day's work, speaks well of their commitment to constantly improve the quality of research at the Air Force Institute of Technology.

We also thank our wives, Sunny and Lynda, who understood our many hours with the computers and reminded us that there is life outside the research environment, and helped us maintain a proper balance, especially when learning how to dance. The dance lessons themselves reinforced our research as we tried to minimize the ratio of toes stepped on divided by the total number of toes!

Ken thanks his little boy, Evan, for bringing lots of smiles and hugs into the study, especially when he wanted to color the dry-erase board or go get "tandy" from Carriage Hill Farm. Ken also thanks Chuck for being the *Microsoft® Word for Windows™* [1] expert and for diligently working hard when Ken was "unavailable."

Chuck thanks his daughter, Charlene, for putting both her driving lessons and her social life on hold while Dad was locked away in his research dungeon. Lastly, Chuck thanks his research partner, Ken, for not only knowing how to thoroughly research a topic, but for also knowing what "parallel verb structures" and "infinitive forms" are.

Ken Hamner

Chuck La Fleur

[1] *Microsoft* is a registered trademark and *Windows* is a trademark of Microsoft Corporation.

Table of Contents

Page

Acknowledgments ii

List of Tables vi

List of Figures vii

Abstract viii

I. Introduction 1-1
- Background 1-1
- Problem Statement 1-3
- Research Questions 1-5
- Scope 1-6
- Assumptions 1-7
- Key Terms 1-7
- Summary 1-8

II. Literature Review 2-1
- Introduction 2-1
- Why Measure? 2-1
- What to Measure 2-2
- Characteristics of Good Metrics 2-3
  - Timely 2-4
  - Drives the Appropriate Behavior 2-4
  - Repeatable and Shows a Trend 2-4
  - Encompasses a Controllable Activity 2-5
  - Feasible to Obtain Data Regularly 2-5
  - Meaningful to the Customer 2-5
  - Formulated at Critical Points 2-6
  - Simple, Understandable, and Unambiguously Defined 2-6
  - Meeting Goals and Objectives 2-6
- Metric Development 2-6
  - The AFSC Method 2-7
  - Basili & Rombach Paradigm 2-9
  - Hayes & Miller Process 2-10
  - The Kinlaw Process Improvement Model 2-11
  - The OMB Generic Method 2-14
  - Extension to the OMB Generic Method 2-16
  - OMB Questionnaire & Checklist Approach 2-16
  - OMB Cost of Quality Failure Approach 2-17
  - Thamhain Model 2-17

Summary 2-18

III. Methodology 3-1
Introduction 3-1
Research Design 3-1
Organizational Backgrounds 3-4
Initial Screening 3-4
Secondary Screening 3-5
Instrument Development 3-8
Survey of Experience 3-9
Metric Development Method Questionnaires 3-10
Procedures 3-11
Data Collection 3-11
Data Analysis 3-11
Summary 3-12

IV. Results and Analysis 4-1
Introduction 4-1
Methods Found in the Literature 4-2
Initial Screening 4-2
The AFSC Method 4-3
Basili & Rombach Paradigm 4-4
Hayes & Miller Process 4-4
The Kinlaw Process Improvement Model 4-4
The OMB Generic Method 4-4
Extension to the OMB Generic Method 4-4
OMB Questionnaire & Checklist Approach 4-5
OMB Cost of Quality Failure Approach 4-5
Thamhain Model 4-5
Initial Screening Summary 4-5
Secondary Screening 4-6
The AFSC Method 4-6
The Kinlaw Process Improvement Model 4-7
The OMB Generic Method 4-8
Secondary Screening Summary 4-8
Metric Development Exercise 4-9
Exercise Participants 4-9
Phase I. Identify the Customers 4-11
Phase II. Define the System 4-11
Phase III. Develop Quality Measures 4-12
Analysis of Results 4-16
Participant Responses 4-17
Researchers' Observations 4-19
Methods Found in Use at ASC 4-20
The AFSC Method 4-20

Page

Participant Experience Levels 4-21
Participant Responses 4-23
Researchers' Observations 4-24
The ASC/SD Method 4-26
Participant Experience Levels 4-26
Participant Responses 4-28
Researchers' Observations 4-29
The ASC/VC Method 4-30
Participant Experience Levels 4-31
Participant Responses 4-33
Researchers' Observations 4-34
Metric Awareness 4-36
Summary 4-37

V. Conclusions & Recommendations 5-1
Introduction 5-1
Conclusions 5-1
Measurement 5-1
Measurement and Metrics 5-2
Metric Development Methods 5-3
Metrics Development 5-3
The OMB Generic Method 5-6
Proposed Metric Development Method 5-7
Recommendations for Metric Development 5-10
Recommendations for Future Research 5-11
Summary 5-12

Appendix A. Survey Of Experience A-1

Appendix B. Sample Metric Development Method Questionnaire B-1

References Bib-1

Vitae Vita-1

List of Tables

Page

Table 1-1. Comparing Typical Process Attributes 1-2

Table 2-1. Sample Service Organization Metrics 2-2

Table 3-1. Initial Screening Criteria 3-5

Table 3-2. Secondary Screening Criteria 3-7

Table 3-3. Secondary Screening Possible Outcomes 3-8

Table 4-1. Initial Screening Results 4-3

Table 4-2. Secondary Screening Results 4-7

Table 4-3. Phase I Results 4-11

Table 4-4. Phase II Results 4-14

Table 4-5. Phase III Results 4-15

Table 4-6. Participant Assessment of the OMB Generic Method 4-19

Table 4-7. Participant Assessment of the AFSC Method 4-25

Table 4-8. Participant Assessment of the ASC/SD Method 4-29

Table 4-9. Participant Assessment of the ASC/VC Method 4-35

List of Figures

Page

Figure 3-1. Research Design 3-3

Figure 3-2. A Sample Likert Scale 3-10

Figure 4-1. The OMB Generic Method 4-10

Figure 4-2. The Faculty Recruitment Process 4-13

Figure 4-3. The AFSC Method 4-22

Figure 4-4. The ASC/SD Method 4-27

Figure 4-5. The ASC/VC Method 4-32

Figure 5-1. The Proposed Metric Development Method 5-9

Abstract

Nonmanufacturing organizations are being challenged to provide high-quality products and services to their customers, with an emphasis on continuous process improvement. Measures of performance, referred to as metrics, can be used to foster process improvement. The application of performance measurement to nonmanufacturing processes can be very difficult.

This research explored methods used to develop metrics in nonmanufacturing organizations. Several methods were formally defined in the literature, and the researchers used a two-step screening process to determine the OMB Generic Method was most likely to produce high-quality metrics. The OMB Generic Method was then used to develop metrics in a nonmanufacturing environment. A few other metric development methods were found in use at nonmanufacturing organizations.

The researchers interviewed participants in metric development efforts to determine their satisfaction and to have them identify the strengths and weaknesses of, and recommended improvements to, the metric development methods used. Analysis of participants' responses allowed the researchers to identify the key components of a sound metric development method. Those components were incorporated into a proposed metric development method that was based on the OMB Generic Method, and should be more likely to produce high-quality metrics that will result in continuous process improvement.

AN EXPLORATORY SURVEY OF METHODS USED TO DEVELOP MEASURES OF PERFORMANCE

I. Introduction

Background

Shortly after President Ronald Reagan issued Executive Order 12637, which called for the Federal Quality and Productivity Improvement Program (Di Lorenzo, 1991:4-1), the Office of Management and Budget (OMB) observed that:

> Federal agencies across government are being challenged to achieve the goal of the Federal Quality and Productivity Improvement effort: to provide high-quality, error-free, and timely products and services to the American public that are responsive to customer needs and make the most effective use of taxpayer dollars. The emphasis is on continuous improvement of operating processes to achieve better products and services and thus, attain greater customer satisfaction.
>
> 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. Measures enable managers to know how close they are to their targets and how to make the right decisions for improving work processes. In short, measures support improvement. This is their key purpose. (OMB, 1989:3)

Continuous process improvement has its roots in manufacturing processes. However, the application of continuous improvement techniques to nonmanufacturing

processes in response to President Reagan's challenge can be very difficult (Tenner and DeToro, 1992:36).

> For many areas outside manufacturing, it becomes a challenge to define appropriate measures of quality, productivity, and market responsiveness. Softer measures of quality and productivity have not only become acceptable, but have become every bit as important as the hard ones. Indicative of this development is the fact that 25% of all possible points in the Malcolm Baldridge National Quality Award scoring system are based on these softer measures. (Thamhain, 1991:471)

Manufactured goods can be measured objectively but "service quality is more abstract and elusive" (Evans and Lindsay, 1993:428). Table 1-1 compares typical manufacturing and nonmanufacturing process attributes, and illustrates the difficulty in measuring those attributes in a nonmanufacturing environment.

Table 1-1

Comparing Typical Process Attributes

Attribute	Manufacturing	Nonmanufacturing
Output properties	Tangible	Intangible or tangible
Production and delivery	Separate	Integrated
Customer interface	Focused: sales and marketing	Spread across line employees
Feedback	Through process	Through customer
Organizational focus	Process efficiency	Customer relations
Process ownership	Clearly defined	Multiple
Process boundaries	Defined	Unclear
Process definition	Documented	Unclear
Control points	Defined	None
Quality measures	Established and objective	Subjective
Corrective action	Preventative	Reactive

(Tenner and DeToro, 1992:40)

Nonmanufacturing processes are different in several other ways. Service organizations often handle large volumes of transactions. For example, the Royal Bank

of Canada processes over 5.5 million transactions daily through their 1600 branches and 3500 banking machines. Second, services are consumed as they are created rather than inventoried or distributed. Services are also more labor intensive, whereas manufacturing is more capital intensive. Finally, services are often very time-sensitive and require a higher degree of customization than manufactured goods (Evans and Lindsay, 1993:26-27).

While measuring quality in a service organization is difficult, it is certainly not impossible. In 1990, Federal Express, the overnight delivery company that handles 1.5 million shipments at 1,650 sites in 127 countries every day, became the first service-category company to win the Malcolm Baldridge National Quality Award. One of the keys to their success was the replacement of their old measure of performance with a better way to measure customer satisfaction in a nonmanufacturing arena (Evans and Lindsay, 1993:26, 131).

Problem Statement

Clearly, research needs to be done in the area of implementing quality measures for service or nonmanufacturing organizations. The U.S. Air Force has service organizations that may have experienced difficulty in identifying quality measures of performance, partly because they are nonmanufacturing organizations. Air Force Materiel Command (AFMC) organizations are identifying performance measures for their own use in the form of metrics. AFMC defines a **metric** as "a measurement made over time, which communicates vital information about the quality of a process, activity, or resource" (AFSC, 1991:2-1).[2]

[2]Air Force Systems Command (AFSC) merged with Air Force Logistics Command (AFLC) on July 1, 1992 to form Air Force Materiel Command (AFMC).

A recent research effort (Hayes and Miller, 1992) evaluated metrics in the Aeronautical Systems Center (ASC) at Wright-Patterson Air Force Base in Ohio. The research team requested internal metrics from five System Program Offices (SPOs) recognized by the ASC Total Quality Office as leaders in the quality arena and received over 300 metrics. Hayes and Miller evaluated a set of those metrics and determined that most of them encouraged behaviors which did not lead to continuous improvement (Hayes and Miller, 1992:xi). Additionally, they suggested that this large number of metrics can be detrimental to the program offices. [These internal metrics are in addition to the many metrics the SPOs already track for the Corporate Management Network (CMN).] They noted that the SPOs had developed their metrics with much less guidance or expertise than had been applied to the AFMC and ASC metrics efforts (Hayes and Miller, 1992:1-5) and recommended further research be done to compare candidate metric development approaches (Hayes and Miller, 1992:5-8).

In addition to the problem cited by Hayes and Miller, a growing discontentment has developed in the SPOs because of the many metrics they must track for reporting to higher levels that do not help local workers manage or improve their own processes. Captain David Garofoli, Director of Measurement at the ASC Total Quality Office, estimated that only ten to twenty percent of the CMN metrics are useful to the SPOs, and one office uses only one out of twenty-five that it tracks (Garofoli, 1993).

Considering the above problems and the fact that another research effort was focusing on evaluating the quality of existing metrics, this research team decided to identify current methods used to develop new metrics in anticipation of finding a way to develop high-quality metrics in the first place.

The team first reviewed the literature relating to measures of performance for nonmanufacturing organizations and uncovered several different metric development methods. However, a comparison of the effectiveness of those metric development

methods was not found. Therefore, it is difficult for nonmanufacturing organizations to develop metrics that truly lead to process improvement because those organizations lack a basis for determining which development method to use. There were also no criteria for judging an existing method's value or guidelines for constructing a new method.

Secondly, the team explored methods in use within the government today and identified the benefits and disadvantages of each method. This procedure enabled them to identify through survey and analysis the key components of a sound metric development methodology and recommend a method that was most likely to result in high-quality metrics. For the purpose of this research, a high-quality metric is one that drives the behaviors which will lead to continuous process improvement.

Research Questions

The following questions guided this research:

1. Why measure?
2. What should be measured?
3. What is the difference between a metric and a measurement?
4. What are the characteristics of a good metric?
5. Which metric development methods are currently formally defined?
6. Which metric development methods are currently in use? What are the strengths and weaknesses of those methods? What improvements have the users recommended for those methods?
7. What are the criteria for a good metric development method?
8. Which of the existing methods would most likely produce high-quality metrics?

Scope

Very little research has been done in the area of metrics development for nonmanufacturing organizations, and since previous studies of quality improvement at the Aeronautical Systems Center (ASC) recommend research to investigate metric development, this effort focuses on metric development rather than metric evaluation (Caudle, 1991:5-1; Hayes and Miller, 1992:5-8). Research on metrics development will provide greater benefit to various organizations desiring to begin the activity of measuring to improve their processes. This research was exploratory in nature and sought to identify metric development methods, either formally defined, or recently used in ASC, and to investigate the applicability of those methods to a nonmanufacturing environment.

The research was also limited to exercising only one metric development methodology. This method was determined to be the most likely to produce high-quality metrics based on an evaluation against the criteria discussed in Chapter 3. This method was exercised in a group session facilitated by the research team for the Faculty Qualification and Recruitment Committee of the Graduate School of Logistics and Acquisition Management.

The research focused on developing metrics for processes specifically related to the fields of education and weapon systems acquisition, and may not be generally applicable throughout all nonmanufacturing environments.

Despite these limitations, the results will still benefit nonmanufacturing organizations which are seeking to develop quality metrics since they report 1) the value of using a systematic process to develop metrics, 2) criteria that a potential method should meet, and 3) the best of the current methods to begin with.

Assumptions

A major assumption for this research effort was that a methodical approach to metric development would best ensure that the metrics developed would lead to continuous improvement. A second assumption that affected the research design was the assumption that participants who had used a method within the last twelve months would sufficiently remember details about the metric development effort.

Key Terms

The following terms are defined with regard to measuring quality in nonmanufacturing organizations and form a knowledge base for the rest of this research effort.

Quality is defined by the American National Standards Institute and the American Society for Quality Control as "the totality of features and characteristics of a product or service that bears on its ability to satisfy given needs." This definition implies that the features or characteristics that relate to quality can be identified and form the basis of a measure, and "satisfy needs" implies that the value of the product or service is determined by the customer (Evans and Lindsay, 1993:9). The customers may of course be internal or external, and the expectations may be implied or explicitly stated (Tenner and DeToro, 1992:31).

Measurement is the process of examining certain characteristics of interest, but may not always communicate vital information. A measurement may not lead to an improved process, but a metric always will (AFSC, 1991:1-1).

Metric is defined as "a measurement made over time, which communicates vital information about the quality of a process, activity, or resource" (AFSC, 1991:2-1). A metric is not an end to itself, but rather a means to achieving objectives and

goals through quality improvement by giving understanding to processes and their capabilities (AFSC, 1991:1-2, 2-2). Managers can use metrics to monitor the health of a process and then take management action as necessitated by the metrics to facilitate an improvement in the performance of a process. Metrics can also be used in benchmarking activities as a basis for comparison with industry practices. **Benchmarking** is the search for industry best practices that lead to superior performance (Camp, 1989:12). This research effort focuses on metrics in particular, but encompasses both benchmarking and measurement.

Summary

Government agencies have been challenged to continually improve organizational processes, and quality measures play the key role in that improvement. However, many Air Force organizations have inundated themselves with quality measures, or metrics, some of which do not truly represent process quality and do not lead to continual improvement. Continuous process improvement has been difficult for nonmanufacturing organizations, but there are several methods available to assist those organizations in developing appropriate measures of performance. This research effort explored these methods to determine the key components of a method which should yield the highest quality metrics.

Chapter two provides a review of the literature relating to measurement in general and examines several metric development methods in particular. Chapter three identifies the process used to determine which of the metric development methods found in the literature was most likely to produce high-quality metrics, and describes the survey instruments and procedures used to identify other methods in use within ASC. Chapter four reports the findings of the surveys, the results of the metric development exercise using the method most likely to produce high-quality metrics, and the analysis of the

exploratory effort. Finally chapter five provides conclusions and recommendations for further research.

II. Literature Review

Introduction

As discussed in chapter one, Air Force organizations are identifying quality measures of performance called metrics. A metric is "a measurement made over time, which communicates vital information about the quality of a process, activity, or resource" (AFSC, 1991:2-1). This chapter discusses measurement in general and metrics and metrics development in particular. The chapter first discusses the importance of measurement, and of knowing what to measure. Then, the characteristics of a high-quality metric are discussed, followed by an examination of several formally defined metric development methods.

Before undertaking a major effort such as defining meaningful measures of performance, it is important to first establish the need to measure.

Why Measure?

Deming's fifth point for management transformation is to "improve constantly and forever the system of production and service" (Deming, 1986:49). Juran, Crosby, Feigenbaum, and others recognize this need for continuous improvement (Tenner and DeToro, 1992:23-24). Measurement is so fundamental to this process that:

> If you cannot measure it, you cannot control it. If you cannot control it, you cannot manage it — it's as simple as that. Measurement truly separates a successful improvement process from one that fails. (Talley, 1991:xi)

Tenner lists measurement as the third step in his six steps to process improvement (Tenner and DeToro, 1992:110). In fact, the Office of Management and Budget (OMB) states that measurement is the only way an organization can know that it is attaining the

desired quality and meeting its customers' requirements (OMB, 1989:9). The most difficult part of measurement for many service agencies is deciding what to measure.

What to Measure

One way to prevent process improvement is to measure the wrong things (Tenner and DeToro, 1992:107). This diverts attention from those things that should be measured, and wastes valuable resources measuring the wrong things. What should be measured? In theory, "measure what's important" (Peters, 1987:483). In practice, it is not so easy. Table 2-1 lists sample metrics used in various service organizations.

Table 2-1

Sample Service Organization Metrics

Organization	Quality Measure
Hospital	Lab test accuracy
	Insurance claim accuracy
	On-time delivery of meals and medication
Bank	Check-processing accuracy
Insurance Company	Claims-processing response time
	Billing accuracy
Post office	Sorting accuracy
	Time of delivery
	Percent express mail delivered on time
Ambulance	Response time
Police department	Incidence of crime in a precinct
	Number of traffic citations
Hotel	Proportion of rooms satisfactorily cleaned
	Checkout time
	Number of complaints received
Transportation	Proportion of freight cars correctly routed
	Dollar amount of damage per claim
Auto service	Percent of time work completed as promised
	Number of parts out of stock

(Evans and Lindsay, 1993:566)

Deming states that the customer is the most important person in the process (Deming, 1986:5), and many agencies are following his advice. The handbook for metrics published by the Office of Management and Budget states, "Quality measures should be directed primarily at customers who are the immediate users of an agency's output and should assess how well their expectations are being met" (OMB, 1989:8). This focus on the customer includes those internal to the process. In fact, better meeting the needs of the next internal customer in the process is critical to continuous process improvement. Dennis Kinlaw explains, "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" (Kinlaw, 1992:98).

One approach to determining what to measure is to identify measures used by other organizations. Measures used by other organizations provide helpful examples, but should never be automatically incorporated without first determining their usefulness in the intended environment. Fortunately, there are several key characteristics of meaningful measures which help determine what should be measured.

Characteristics of Good Metrics

The Air Force Materiel Command identifies eight attributes of a good metric (AFSC, 1991:2-1), and OMB suggests five criteria for assessing measures of performance (OMB, 1989:16). Combining the similar concepts yields the following characteristics which are listed below and described in the following paragraphs.

1. It is timely
2. It drives the appropriate behavior
3. It is repeatable and shows a trend
4. It encompasses a controllable activity
5. It is feasible to obtain the data regularly
6. It is accepted as meaningful to the customer
7. It is formulated at a critical point in the process

8. It is simple, understandable, and unambiguously defined
9. It tells how well organizational goals and objectives are met through processes and tasks

Timely. Metrics must be collected in a timely fashion, allowing any improvement efforts to be based on the most current information. Ideally, process improvements would occur in a real-time fashion. Realistically, the more critical the process, the faster the resulting improvement needs to occur.

Drives the Appropriate Behavior. Most important of all, the metric must lead to behavior that results in continuous improvement. For example, if programmer productivity is measured using lines of code produced, the lines of code will surely increase. The lines of code may be inefficient, unnecessary, or even counterproductive, but the measurement falsely indicates a productivity increase. One insurance firm used the percentage of claims paid within two days of receipt as a measure of accuracy in paying claims, and fostered its employees to pay more claims within two days, including claims which should have been disallowed (Kerr, 1975:778). This aspect is so crucial to the quality of the metric that Hayes and Miller dedicated their research primarily to two issues related to this concept: 1) What behavior is this metric likely to drive? and 2) How much will this behavior contribute to continuous improvement? (Hayes and Miller, 1992:3-10). Generally, if the measure does not help improve the process or support an objective, it is a waste of resources and should be avoided.

Repeatable and Shows a Trend. A metric must be repeatable; that is, it must be capable of providing the same result if the process is in the same state, even if the measure is performed by a different person. Also, the metric should reflect progress toward improved quality by showing if quality is getting better or worse (Juran, 1988:77). A one-time measure could show whether a goal had been achieved, but would

be insufficient to determine convergence toward a goal or divergence from the goal. Conversely, a measure over time can give tremendous insight into the health of a process.

Encompasses a Controllable Activity. The metric must be based on some aspect of quality that is under the group's direct control (Kinlaw, 1992:110). Otherwise, it would be wasteful to measure a part of a process without having the authority to change it. "Since the intent is to use the information to verify and make improvements, it is important that the measure reflects any action taken to change the process" (OMB, 1989:16).

Feasible to Obtain Data Regularly. The use of a metric must add value to a process. If the effort to collect the measurement data exceeds the potential benefit from the resulting improvement to the process, then the measurement should be discarded. Also, regarding precision of data, Juran states:

> 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 criteria 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.
>
> To measure arrival and departure time of commercial airlines to the nearest minute is close enough. To extend the precision to the nearest second would mean extra effort in definitions, measurement, and so forth, with doubtful value. (Juran, 1988:78)

Meaningful to the Customer. The customer should be considered when developing and validating any metric an organization uses. In fact, the customer is the one who defines the quality of output for any given process. Therefore, the meaning of the metrics should translate to the needs and expectations of the immediate and ultimate

users by defining internally measurable characteristics which, if improved, would better satisfy the customers (AFSC, 1991:3-3).

Formulated at Critical Points. "The metric must be 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" (OMB, 1989:16).

Simple, Understandable, and Unambiguously Defined. Complex metrics are ineffective. They should be simple enough to be understood by everyone involved in the process, including the customer. "Quality measurement is only effective when it is done in a manner that produces information people can understand and use" (Crosby, 1979:199). The metric must also be based on clear operational definitions which have been agreed to by all interested parties. Unfortunately, "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" (Juran, 1988:76).

Meeting Goals and Objectives. Metrics should be based on the goals of an organization or process, and a metric should clearly indicate whether a goal is being achieved or not. While progress toward a goal is a worthwhile activity, it is a necessary, but not sufficient condition for a metric — a metric is a means to achieving objectives and goals through process improvement (AFSC, 1991:2-2).

Metric Development

Existing metric development methods were examined to determine their suitability for use in nonmanufacturing organizations. There are several methods formally defined or in use today. The best one to use may depend on the structure of the process of

interest and the needs of the organization (OMB, 1989:10). An examination of each of the methods currently available follows.

The AFSC Method (AFSC, 1991:3-1 to 3-5). *The Metrics Handbook* outlines a ten-step process for developing a metric package, which includes the operational definition of the metric, the actual measurement data, and the presentation of the metric. There is currently a draft AFMC Pamphlet 74-9 dated 26 March 1993 that has a few updates to the original handbook. However, the steps in the metric development method are still the same.

Step I. *Identify Your Purpose*

> It is important to first align your purpose with your organization's mission, goals, and objectives. These should focus on meeting customer needs and serve as a foundation for accomplishing and sustaining continuous, measurable improvement.

Step II. *Develop Your Operational Definition Starting With Your Customer*

> Define the who, what, when, why and how of this metric in sufficient detail to permit consistent, repeatable and valid measurement to take place. The operational definition starts with an understanding of your customers' expectations. You then "operationalize" the expectation by defining the characteristic [sic] of the product, service, or process which are internally measurable and which, if improved, would better satisfy your customers' expectations. This is actually an iterative process involving Steps II-VII. This is the first element of your metric package.

Step III. *Identify And Examine Existing Measurement Systems*

> Once you have established the link to objectives and goals, it is essential to determine if existing metrics or other measurement systems exist that satisfy your requirements. Don't "reinvent the wheel". Use existing process measurements when they exist.

Step IV. *Generate New Metrics If Existing Metrics Are Inadequate*

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 drive the final outcome and are the key to making process improvements. The assumption is: if you monitor and improve process performance, the quality of the products and services will improve.

Step V. *Rate Your Metric Against The "Eight Attributes of a Good Metric"*

Refer to the attributes listed in Chapter Two [of *The Metrics Handbook*, discussed here beginning on page 2-3.] If you feel your metric sufficiently satisfies these criteria for a good metric, go on to Step VI. If not, return to step II and correct the deficiencies.

Step VI. *Select Appropriate Measurement Tools*

Select the proper tool for analyzing and displaying your data. The "Eight Basic Tools" identified in Figure 4-1 [of *The Metrics Handbook*] are those most commonly used. Other statistical and non-statistical tools may be more appropriate for your application. Use whatever you feel is best. The tools will be discussed further in Chapter Four and Appendix C [of *The Metrics Handbook*].

Step VII. *Baseline Your Process*

Start acquiring metric data. This serves as a baseline for determining the capability of your process. Ask if the data is accumulated over time and adequately measures the important characteristics of your process. If the answer is uncertain, examine other possibilities. And if you change your metric, remember to coordinate it with your customer again.

Step VIII. *Collect And Analyze Metrics Data Over Time*

Continue aggregating metric data over time. Examine trends. Special and/or common cause effects on the data should be investigated and assigned. Compare the data to interim performance levels. This is the second element of your metric package.

Step IX. *Finalize The Metric Presentation*

> Based on the results of the previous steps, you are finally ready to present the metric externally. The descriptor will provide enough information to communicate the appropriate details of your metric to your customer. Determine the appropriate level of detail through discussion with the customer. This information should be an abbreviation of the key elements of the operational definition. The graphic presentation clearly and concisely communicates how you are performing. This is the third element of your metric package.

Step X. *Initiate Process Improvement Activities*

> Initiate process improvement activities in conjunction with the key process owners. Chapter Six [of *The Metrics Handbook*] provides more guidance in this area. Once you have implemented improvements, the process above may start over or it may pick up again at almost any step. Remember, continuous improvement requires continuous effort. THIS STEP IN THE DEVELOPMENT PROCESS IS THE MOST CRITICAL FOR YOUR IMPROVEMENT EFFORTS TO BECOME A REALITY. Remember that metrics are just a means to an end! That end is continuous process improvement.

The AFSC method is a step-by-step procedure for developing organizational metric packages, and it focuses on the external customers only.

Basili & Rombach Paradigm (Basili and Rombach, 1987:350-351). Basili and Rombach provide a goal/question/metric paradigm within a framework for improving the software development process that may be generic enough to apply to other nonmanufacturing activities. This approach provides for generating goals and deriving quantifiable questions and metrics.

1. *Generate a set of goals based upon the needs of the organization.*

> The first step of the process is to determine what it is you want to improve. This focuses the work to be done and allows a framework for determining whether or not you have accomplished what you set out to do. Sample goals might consist of such issues as on [sic] how to improve the set of methods and tools to be used in a project with

respect to high quality products, customer satisfaction, productivity, usability, or that the product contains the needed functionality.

2. *Derive a set of questions of interest or hypotheses which quantify those goals.*

   The goals must now be formalized by making them quantifiable. This is the most difficult step in the process because it often requires the interpretation of fuzzy terms like quality or productivity within the context of the development environment. These questions define the goals of step 1. The aim is to satisfy the intuitive notion of the goal as completely and consistently as possible.

3. *Develop a set of metrics and distributions which provide the information needed to answer the questions of interest.*

   In this step, the actual data needed to answer the questions are identified and associated with each of the questions. However, the identification of the data categories is not always so easy. Sometimes new metrics or data distributions must be defined. [Example removed for clarity.] As the data items are identified, thought should be given to how valid the data item will be with respect to accuracy and how well it captures the specific question.

The authors also differentiate between questions that focus on the process quality and questions which focus on product quality.

Hayes & Miller Process (Hayes and Miller, 1992:5-6). Hayes and Miller recommend a metric development process based on their weapon systems acquisition-related metric evaluation research.

1. Select an acquisition objective.
2. Brainstorm the positive behaviors that would help meet this objective.
3. Brainstorm metrics that will drive these behaviors.
4. Evaluate the metrics via the cloud chart method described.

While this method focuses on driving the right behavior, it does not expand on how to perform each step. It does not specifically include the customer, and it is output rather than process-oriented.

The Kinlaw Process Improvement Model (Kinlaw, 1992:106-111). Dennis Kinlaw describes a metric development method within his framework for improving a process. The basic process improvement framework is to:

Step 1. Understand The Opportunity Or Problem.
Step 2. Define The Specific Improvement Target
Step 3. Design Strategies To Reach The Target
Step 4. Design The Data Links
Step 5. Design The Response Process To Use Data From The Data Links
Step 6. Determine How The Project Will Be Managed

Kinlaw's metric development process is part of Step 1 in the process improvement framework, but before arriving at the metric development process, answering four key questions will increase the understanding of the opportunity or problem.

Question 1. Who are our customers?
Question 2. What is the measured quality of our services and products?
Question 3. What are our customers' perceptions of our output, i.e., our services and products?
Question 4. What are our customers' expectations of our output?

Metrics are developed by determining the measured quality of an organization's services and products, i.e., by answering question two. Kinlaw recommends the following steps:

1. Identify the output

    a. *Develop clarity about the meaning of the output.*

        Outputs are all the services and products that the team delivers for the use of an internal or external customer. It is useful for a

team to refer to its list of customers and to think about each of these customers as it produces its list of outputs. All information that is developed by the team or that the team is expected to use should be recorded and displayed on charts.

b. *Develop clarity about measures of output as ratios.*

One thing is always compared to another, i.e., production unit to production time or service unit to errors.

c. *Develop by brainstorming, a list of as many outputs as possible.*

The team should peruse its lists of customers and, for each customer, develop two lists of outputs, one for services and one for products.

d. *Review all outputs and ensure that they meet the following criteria*:

1) They are phrased in concrete terms so there is no question about their exact meaning
2) They represent something for which the team has total or shared responsibility
3) The team can directly influence some aspect of the output's quality

Note: Outputs for which the team shares responsibility should be set aside for joint improvement projects later.

2. Establish the relative importance of the output

a. *Develop criteria and select the outputs for which measures will be developed first.*

Criteria can be based on the importance of the customer who is using the output, the relative dollar value of the output, the relative visibility of the output, the relative importance of the output to upper management, the relative amount of the output, etc.

3. Develop measures for assessing the quality of the output

a. *Transfer the first output for which measures will be developed to the top of a separate page of chart paper and draw a line under it.*

This is the numerator of the measurement ratio that is being developed.

b. *Establish as a team the specific kind of measure that will be developed.*

Two types of measures I [sic] recommend are attribute measures and cost measures.

c. *Develop, by brainstorming, a list of the sources of data that the team will use to measure the output.*

If the measure is an attribute measure, the sources of data will be types of errors or failures. If the team is developing a cost measure, they will be things such as travel, planning, direct labor, hardware, rework, etc. Record the list under the numerator (i.e. the output) to be measured. This provides the team with a picture of the relationships that it is trying to establish in the ratios that it is building.

d. *Select from the data sources listed in the denominator the ones that will be tracked.*

First establish selection criteria such as availability of data, times spent in acquiring data, amount of information that the data will provide for potential improvement, etc.

e. *Review the ratio and refine it if necessary.*

4. Use the measures to track and monitor improvement in the quality of the output.

A team using Kinlaw's method would then answer the other three key questions (listed here on page 2-11) leading to understanding the problem, and the team would then complete the remaining steps of the entire process improvement framework (also listed here on page 2-11). Several examples of measurement ratios are provided, as well as more detail on the remaining steps of the process improvement framework (Kinlaw, 1992).

The OMB Generic Method (OMB, 1989:10-16). OMB's Generic Method is an approach that looks at every step of a process and translates attributes of quality, as defined by the customers, into indicators or measures of quality. This method includes worksheets and examples that accompany each step.

1. *Identify all customers of the program's outputs -- products and services -- and those customers' requirements and expectations.*

   One needs first to answer the questions: "Who are my customers?" "For whom are we producing a final output (service or product)?" When this is answered, the next step is to define those customers' needs and expectations by asking them the question, "what are the quality attributes (e.g. accuracy, consistency, clarity, responsiveness) that must be satisfied to meet your expectations and requirements?" This procedure should be followed for each output. [A worksheet is provided to accomplish this step. An example is also provided, followed by more guidance.]

2. *Define the entire work process that provides the product/service.*

   After customers are identified and their expectations are determined, the team developing measures then needs to consider the total work process or program function that produces the agency output. This is important in order to find out if the process enables the agency to fulfill customer expectations.

   First, define the function's objective or purpose. Next, identify the first and last steps taken to produce the output in order to determine the parameters of measurement. The total function will most likely, but not necessarily, involve more than one work activity. [A worksheet is provided to accomplish this and the next step. The previous example is continued, followed by more guidance.]

3. *Define the value-adding activities and outputs that comprise the system.*

   Identify each step in the system where "value is added" and an *intermediate output* is produced. This step should lead to weeding out steps that do not add value to the process, such as extraneous procedures, non-essential requirements, and time-consuming approval points. Just as customers are identified for the program's final output, identify the customers or users of the intermediate outputs. Next,

determine the needs and expectations of each intermediate customer. At each step, the quality of the earlier procedures affects the ability to perform the current step with quality. [A worksheet is provided to accomplish this and the previous step. The running example is continued, followed by more guidance.]

4. *Develop quality measures or indicators.*

Each interaction or step indicated above represents a critical point at which value is added to the output for the next user/customer until the final output is produced or delivered. These steps, therefore, become important checkpoints for measuring quality. Consider then, "What should I look at to gauge how well the process is producing intermediate outputs that meet each customer's needs and expectations?" For each step, determine the key deviations that produce problems or variations in meeting customer needs and expectations. Ask the question: "What is the source of that variability?" Answers to these questions indicate why quality is or is not achieved at each critical point. This descriptive information then needs to be quantified by putting it into a ratio format. [A worksheet is provided to accomplish this and the next step. The previous example is continued, and several example metrics in ratio format are also provided.]

5. *Assess quality measures.*

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?

The OMB Generic Method starts with process goals rather than organizational goals, and focuses on both the customer and the intermediate steps of the process (rather than just final results). Guidance is provided on how to perform each step.

Extension to the OMB Generic Method. The extension adds a sixth step after the metrics have been produced. Step six is the development of a matrix (seven steps for this) used to weight and aggregate the measures of multiple characteristics into a single quality index for the entire process of interest (OMB, 1989:16-18).

This can be an effective way to monitor the progress of an entire organization. For example, Federal Express won the 1990 Malcolm Baldridge National Quality Award using a system like this that aggregates service quality indicators from all over the world to give them a single customer satisfaction rating, which is televised daily to every Federal Express location, keeping "everyone moving toward the target of 100 percent customer satisfaction" (Galagan, 1991:27-28). However, this would not be an effective method for developing individual metrics.

OMB Questionnaire & Checklist Approach (OMB, 1989:18-23). The OMB Questionnaire & Checklist Approach sets up a structure that measures quality by identifying significant areas of activity in a particular function.

1. Determine the significant areas of activity where quality is important to the customers (these may be determined by legal or regulatory requirements as well as customer feedback).

2. Ask the question: "What do I look at in each area to gauge how well the process meets customer needs or expectations?" With this information, formulate a series of questions in each area that require a YES or NO answer. Sources for questions could be regulations or policy requirements, standards used by an appropriate accreditation

board, a supervisor's or manager's expectations, and customer feedback.

3. Have a trained person (or supervisor) check each activity on a random basis several times during the year and respond to each question with a yes or no response. Calculate the percentage satisfactory responses (yes) in each area to yield a raw score.

4. Rank each area of activity by its importance relative to the other areas. This is done using some agreed-upon criteria — it may be that one area [examples removed for clarity] clearly adds greater value to the final product than another. Weight these areas appropriately by assigning a percentage figure which reflects their relative value. Weights should add up to 100%.

5. Multiply the raw score by the relative value of each area to arrive at a quality index for each area of activity.

6. Add together all area indexes to get one total quality index. This figure will be a percentage indicating quality achieved by the function [example removed for clarity] out of a possible 100% score. While this step allows a net aggregate quality index to be calculated, it is important to evaluate the quality indexes for each area. These indexes give information about the critical points in the total process and point to areas in need of improvement.

In this case the checklist is the measurement tool. Checklists may be better at measuring conformance rather than driving continuous process improvement.

OMB Cost of Quality Failure Approach. This approach combines a sequential look across the work processes with the use of checklists and transforms the quality failures into an actual dollar or cost figure. Once the total cost of quality failures (both internal and external) is determined, the organization strives to drive that cost down to zero (OMB, 1989:22, 23).

Thamhain Model (Thamhain, 1991:474). Thamhain recommends a method which employs a group process to answer the following questions that will lead the group to the metrics it needs:

1. What is critical to business success?
2. What support areas do we depend on?
3. What materials, processes, services are critical?
4. What organizational linkages are critical?
5. What are our deliverables?
6. Who are our customers/clients?
7. What affects our costs, quality, timing?
8. Where do we experience operational problems?
9. What management processes affect our performance?

Summary

Before reviewing the literature for formally defined metric development methods, the research effort focused on laying a proper foundation for continuous process improvement. This research examined why measurement is an important part of improvement, and what types of things should be measured. The characteristics of high-quality metrics were then presented, followed by descriptions of the metric development methods defined in the literature. Several methods were found, ranging from a method described in a single paragraph to one described in an entire book. A scientific approach for comparing the effectiveness of those metric development methods was not found. Criteria for judging an existing method's value and guidelines for constructing new metric development methods were also not found. The lack of comparative analyses and further lack of analysis tools with which to perform a comparison provides fertile ground for the remaining research effort, which will identify the formally defined metric development method most likely to produce high-quality metrics.

III. Methodology

Introduction

This chapter describes the method used to conduct the research effort. The overall research design is presented, followed by a section detailing how the instruments were developed. The procedures used for data collection and analysis are then presented.

Research Design

This exploratory effort was a field study intended to investigate current practices in metric development for nonmanufacturing organizations. A field study is more rigorous than a case study in that it attempts to "gain a more balanced view before drawing conclusions" (Gibson and others, 1991: 698, 699). The researchers used two sources to identify metric development methods for inclusion in this research and interviewed participants from actual metric development efforts in order to gain a complete view of current practices before drawing conclusions. The first source was the literature relating to measures of performance for nonmanufacturing organizations. Nine metric development methods formally defined in the literature were included in this research. The researchers then used a two-step screening process to determine which of those documented methods would most likely produce high-quality metrics. The metric development method which emerged from the screening process was then used by faculty members of the Air Force Institute of Technology (AFIT) Graduate School of Logistics and Acquisition Management. This provided the researchers the opportunity to interview persons who participated in the use of the metric development method considered most likely to produce high-quality metrics.

The second source of metric development methods was the Aeronautical Systems Center (ASC). ASC[1], a product center of the Air Force Materiel Command (AFMC), was particularly suitable for this field study because it has made an active commitment toward implementing Total Quality principles and has won the Federal Quality Institute's "Quality Improvement Prototype" Award (ASD, 1992:1). This award recognizes federal organizations that have improved their products and services through the measurement and analysis of the efficiency, quality, and timeliness of associated processes (FQI, 1990:21). ASC strives to continually improve its products and services with the help of metrics (ASD, 1992:1).

A few organizations within ASC were identified by the Total Quality Office as having undergone metric development efforts within the past twelve months. The researchers limited the time frame to twelve months to enhance the interview participants' ability to recall the details of their metric development experiences. Three organizations that had undergone metric development efforts were included in this research.

The researchers interviewed participants in those efforts to determine which, if any, metric development methods were employed. The research team first interviewed the person(s) most knowledgeable of the metric development effort to define the actual method used. Then the team created a flowchart of the method and verified the accuracy of the flowchart with the person(s) originally interviewed. Finally, other persons who participated in the metric development efforts were identified and contacted to participate in the next step of interviews as shown in Figure 3-1.

Regardless of the method's source (literature or ASC), participants in metric development methodologies were interviewed using a survey instrument designed to

---

[1]The Aeronautical Systems Division (ASD) became the Aeronautical Systems Center (ASC) on July 1, 1992.

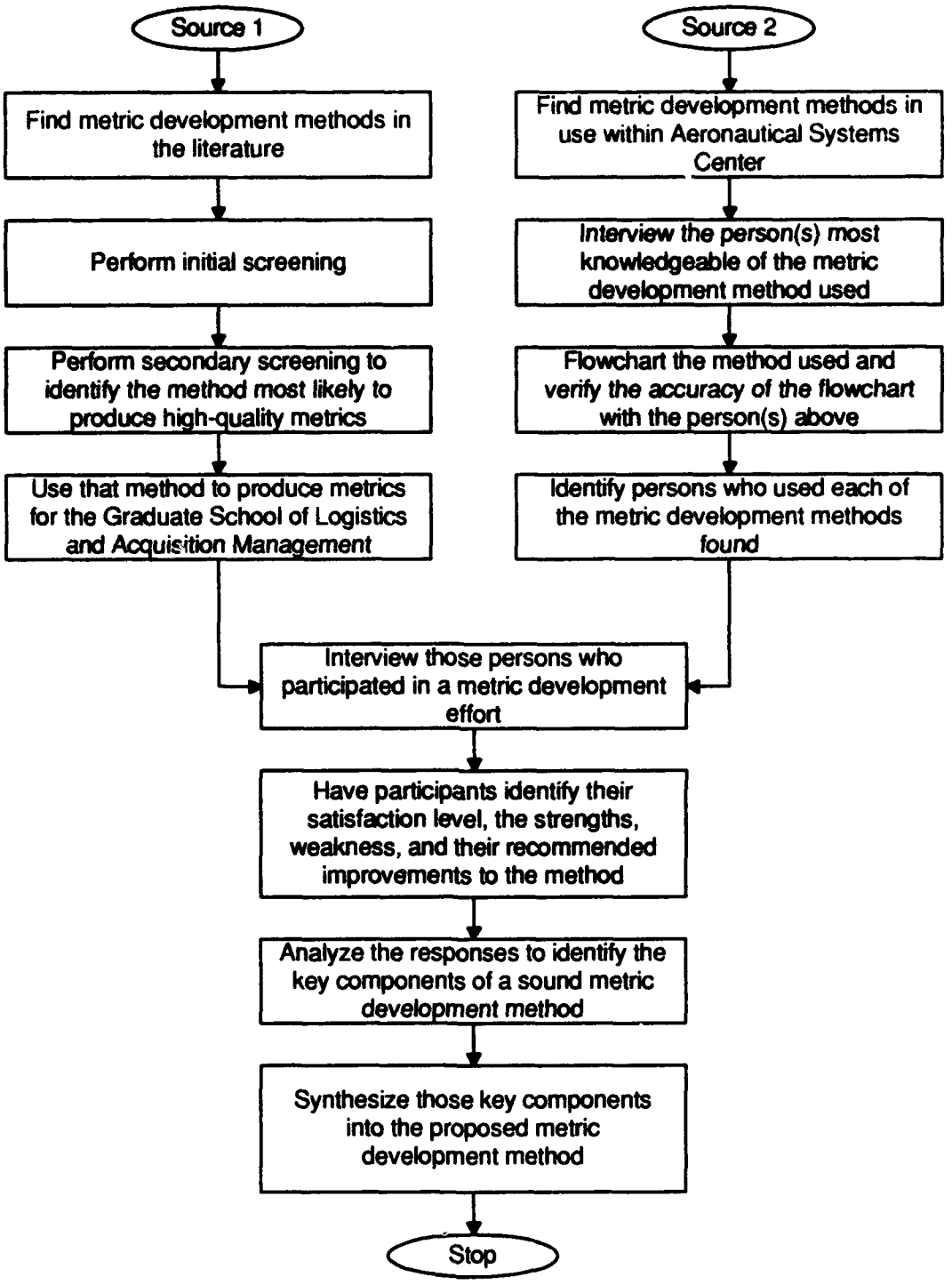


Figure 3-1. Research Design

measure participant satisfaction with the method used, and to allow participants to identify strengths, weaknesses, and recommended improvements to that method. The research design was implemented according to the flowchart shown in Figure 3-1.

Organizational Backgrounds. The exploration of current practices in metrics development included participants from four organizations: The Directorate of Acquisition Logistics (ASC/AL), the Aircraft SPO (ASC/SD), the Advanced Cruise Missile SPO (ASC/VC), and the AFIT Graduate School of Logistics and Acquisition Management. Each had particular reasons for developing metrics and processes they were interested in improving. These will be covered in more detail in chapter four.

ASC/AL took the lead in developing certain administrative metrics that were common to all the functional areas within ASC (contracting, engineering, financial management, program management, and logistics). ASC/SD, which is a "basket SPO" with twenty-four programs managed by Integrated Product Teams (IPTs), developed measures of cost, schedule, performance, and supportability that could be standardized between the teams. ASC/VC developed metrics for many IPTs to be used with their stand-alone management system called the Acquisition Program Tracking System (APTS). Finally, the Faculty Qualification and Recruitment Committee developed metrics to improve the school's process of selecting qualified candidates for Ph.D. sponsorship and faculty training. There were sixteen participants interviewed in the four organizations.

Initial Screening. During the initial screening, the research team rated each of the metric development methods against the criteria listed in Table 3-1. The researchers openly discussed the merits of each method against the stated criteria until the researchers agreed upon the rating assigned. A metric development method needed to

satisfy all three criteria in Table 3-1 to be eligible for inclusion in the next level of screening.

Table 3-1

Initial Screening Criteria

Criteria	Yes	No
1. Will the method result in continuous process improvement?		
2. Is there sufficient detail provided in the literature to actually execute the method?		
3. Does the method focus on individual processes, rather than an entire system?		

First, the method needed to potentially result in process improvement rather than simply report the status of the process. Fostering continuous improvement of the process is what separates a metric from a measure. This is analogous to a gauge on a dashboard instead of an indicator light.

There had to be sufficient detail in the method for the research team to exercise it in a real-world environment. Vague language or insufficient detail might lead to misinterpretation or problems in conducting each step of the method. Also, methods based on lists of questions did not pass this step if they did not include a specific way to transition from the question's answer to a quality measure.

Finally, the method needed to encourage the developer(s) to focus on an individual process rather than the system as a whole. Process improvement is most likely to occur when quality measures are developed from the bottom up, not from the top down.

Secondary Screening. This level of screening was performed on those metric development methods which survived the initial screening by a panel consisting of

the research team and the research advisors. The metric development methods were rated against the criteria listed below. The purpose of the secondary screening was to identify the method most likely to produce high-quality metrics. A summary of the secondary screening criteria is included in Table 3-2.

Understandable. Each step in the method needed to be clear enough that participants will understand exactly what should be accomplished in that step and why it should be accomplished.

Executable. Each step needed to reflect a task that could actually be accomplished within reasonable constraints (time, data, skill levels, etc.).

Efficient. Each step needed to add value or provide a clear benefit to the metric development process without being redundant or creating intermediate products that were not used.

Logical. The method needed to logically flow from one step to the next with the output of one step contributing to the execution of the next one.

Flexible. The method needed to be flexible enough to be used in different settings. Some developers may be able to exercise the method at one sitting for example. Others may need to split the task into separate sessions.

Sufficient. The method needed to include everything necessary to yield high-quality metrics and not omit key steps, instructions, definitions, or examples.

Process-Oriented. The method needed to encourage the developers to focus on the process rather than the results or outputs. The outputs can only be improved by improving the process.

Assessable Results. The method needed to provide some way to assess the value of the metrics that were produced.

Helpful Tools. Any tools (examples, worksheets, etc.) provided to aid the developers in performing each step needed to be helpful.

Customer-Oriented. The method needed to encourage the developers to focus on the customers who were both internal and external to the process in question.

Table 3-2

Secondary Screening Criteria

Criteria	Rating			
	Blue	Green	Yellow	Red
	Quite a bit	Moderately	Barely	Not at all
1. How understandable were the steps? (Was the purpose of each step clear?)				
2. How executable were the steps? (Could you actually perform them?)				
3. How efficient was the method? (Did each step add value?)				
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)				
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)				
6. How sufficient was the method? (Did the guidance list all the steps necessary to develop a metric?)				
7. How process-oriented was the method? (Did the guidance force you to focus on a process needing improvement?)				
8. How assessable was the result? (Was a mechanism provided to assess the quality of the metric produced?)				
9. How helpful were the examples provided for each step of the method?				
10. How well did the guidance force you to focus on the customer of the product or service provided?				

Each participant in this screening cast a secret ballot assigning a rating for each criteria. Each participant's vote carried the same weight. With four participants and four ratings available, there were 4^4 possible outcomes. Those 256 possible outcomes were placed into four categories, with each category having a specific action required. Table 3-3 lists the categories for all possible outcomes, and the action resulting from that outcome. This secret ballot process continued for each criteria listed in Table 3-2, until a single method had been completely assessed. The balloting then continued for the next method under consideration.

Table 3-3

Secondary Screening Possible Outcomes

Outcome	Examples	Action
1. Four participants assign the same rating.	Green, Green, Green, Green	Accept the rating.
2. Three participants assign the same rating, the fourth participant assigns a rating which differs by a single level.	Green, Green, Blue, Green	Accept the majority rating.
3. Two participants assign one rating, the other two participants assign a rating which differs by a single level.	Green, Green, Blue, Blue	Participants discuss the rationale for their assessment. Another secret ballot occurs.
4. Participants assign ratings which differ by more than one level.	Green, Green, Red, Green Green, Blue, Yellow, Green	Participants discuss the rationale for their assessment. Another secret ballot occurs.

Instrument Development

Survey respondents and survey instruments can both be sources of error when conducting measurements (Emory and Cooper, 1991:178). The researchers had little

control over the respondents as a source of error, but some control over the instrument as a source of error was possible. Instrument error was reduced by pre-testing each of these instruments, allowing the researchers to refine the instruments before actual use and increasing the validity of the measuring tool (Emory and Cooper, 1991:180, 422). This research effort required a survey of experience, and a questionnaire for each of the metric development methods used to produce metrics.

Survey of Experience. The survey of experience (See Appendix A) was used to collect data for the purpose of characterizing the participants' backgrounds, experience levels, and familiarity with the area of quality in general, and metric development in particular. This survey collected demographic information (name, age, job title, time spent in that job) and included background questions regarding the level of general education received, and the level of education and or training with regard to the subject of quality. Respondents were asked to select from a list of candidate authors of books or papers on the subject of quality, and then to annotate which of those candidate authors had written work(s) the respondent had actually read. The candidate list of 21 individuals included 4 individuals (Nancy Brady, Sonja A. Seefeldt, Carl J. Walt, and Lynda Yates) who had not authored material on the subject of quality. Those four individuals were included to assist in identifying respondents who may have been trying to guess which candidates had authored the material in question. The 17 remaining individuals constituted a representative list of authors on the subject of quality rather than an exhaustive list. Finally, the respondent was asked to define the term metric as it relates to quality, and to identify any metric development methods with which the respondent was familiar. Pre-testing of this instrument did not identify any problem areas with the instrument.

Metric Development Method Questionnaires. The metric development method questionnaires (See Appendix B) were used to measure participant satisfaction with a metric development method that was used by that participant. Measuring abstract items such as opinions and attitudes concerning satisfaction is frequently done through the use of ordinal scales, such as the Likert scale. This scale required respondents to grade their degree of agreement with a statement on a multi-point scale (Emory and Cooper, 1991:220). Figure 3-2 provides a sample of a Likert scale used to measure a person's attitude toward research.

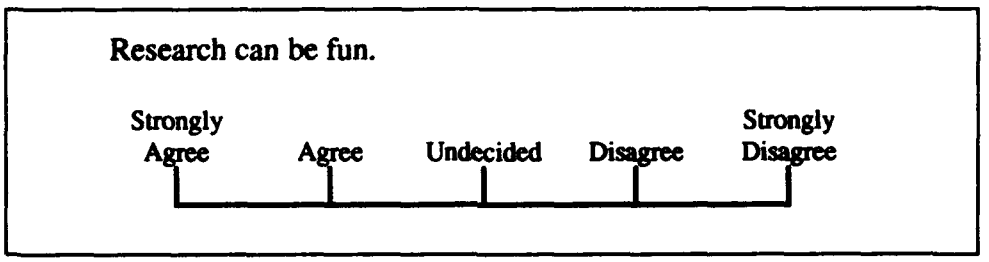


Figure 3-2. A Sample Likert Scale

The researchers opted for a four-point ordinal scale for those items requiring respondents to grade their degree of agreement with a statement. Also, the researchers included a color-rating scheme (Blue, Green, Yellow, Red) similar to the scheme used in various acquisition situations such as source selections and program reviews. Using a scheme respondents were familiar with may have assisted them in choosing the answer most closely matching their degree of agreement.

The first page of each metric development method questionnaire contained a flowchart of that method as drawn by the researchers. The respondents were first asked to verify the depicted flowchart represented the process they actually used.

The remaining three pages of each questionnaire were identical, regardless of the method used. Those pages were designed to measure participant satisfaction with the

method used, and to allow participants to identify strengths, weaknesses, and recommended improvements to that method.

Pre-testing this instrument uncovered a flaw in the original design. The secondary screening criteria listed in Table 3-2 were initially included as page two of each questionnaire. After assessing the method using page two, the respondents were asked more detailed questions on pages three and four of the original instrument. Those subsequent questions tended to cause respondents to give the matter more thought, and then give answers conflicting with their original assessment. Therefore, the researchers moved the assessment criteria from page two to page four of the revised questionnaire, allowing respondents to assess the method as the final step in their effort, reducing the potential for inconsistent responses.

Procedures

Data Collection. First, data was collected with a survey designed to characterize the participants' backgrounds, experience levels, and familiarity with the area of quality in general, and metric development in particular. The survey of experience data was collected from any person interviewed as part of this research effort, regardless of whether they had participated in a metric development effort or not. Participant satisfaction data was collected from each person interviewed who actually developed metrics.

Data Analysis. The researchers analyzed participant responses to identify those steps in the metric development method used by each participant which should be included in the method proposed by the researchers. The participant responses included their satisfaction with the metric development method used, and the strengths, weaknesses, and recommended improvements they identified for that method. The final

analysis led to identification of the key components of each method, resulting in a proposed metric development method which should produce higher quality metrics than any of the methods currently used.

Summary

Canvassing both the literature and the field as sources of metric development methods provided the best opportunity to uncover as many methods as possible. Determining which of the methods documented in the literature would most likely produce high-quality metrics allowed the researchers to facilitate the Air Force Institute of Technology's Graduate School of Logistics and Acquisition Management's use of that method to develop actual metrics. Comparing participant satisfaction concerning that documented method with other participants' satisfaction concerning other methods actually used enabled the researchers to identify the key components of each method, resulting in a proposed metric development method which should produce higher quality metrics than any of the methods currently used.

IV. Results and Analysis

Introduction

This chapter presents the results of this research effort. A listing of the metric development methods found in the literature is presented, followed by the results of the two-step screening process used to identify the method most likely to produce high-quality metrics. The Office of Management and Budget (OMB) Generic Method was determined to be most likely to produce high-quality metrics. A flowchart of this method is presented.

The OMB Generic Method was then used in a metric development exercise by the Faculty Qualification and Recruitment Committee of the Air Force Institute of Technology's Graduate School of Logistics and Acquisition Management. The results of this exercise are also presented.

The second source of metric development methods was the Aeronautical Systems Center (ASC). A few organizations within ASC were identified as having recently undergone metric development efforts, and descriptions of those metric development methods are presented. Flowcharts of those methods are also presented.

Participants in metric development efforts were interviewed, regardless of the source, and the results of those interviews are reported. Participant experience levels are presented, as well as the data resulting from surveying participant satisfaction with the metric development methods used. Metric development participants identified strengths and weaknesses of the metric development methods they used, and the participants' recommended improvements to those methods are presented.

Methods Found in the Literature

A search of the literature relating to measures of performance for nonmanufacturing organizations yielded several different metric development methods. Refer to chapter two for descriptions of these methods.

- The AFSC Method
- Basili & Rombach Paradigm
- Hayes & Miller Process
- The Kinlaw Process Improvement Model
- The OMB Generic Method
- Extension to the OMB Generic Method
- OMB Questionnaire & Checklist Approach
- OMB Cost of Quality Failure Approach
- Thamhain Model

Initial Screening

The initial screening rated each of the metric development methods found in the literature against the criteria listed in Table 3-1. This screening was done by the research team only. The researchers openly discussed the merits of each method against the stated criteria until the researchers agreed upon the rating assigned. A metric development method needed to satisfy all three criteria in order to be carried forward in the screening process. Table 4-1 provides a summary of the initial screening results. Those methods removed from further consideration are shown as shaded rows in the table.

Table 4-1

Initial Screening Results

Method	Will the method result in continuous process improvement?	Is there sufficient detail provided in the literature to actually exercise the method?	Does the method focus on individual processes, rather than an entire system?
The AFSC Method	Yes	Yes	Yes
Basili & Rombach Paradigm	No	No	Yes
Hayes & Miller Process	Yes	No	No
Kinlaw Process Improvement Model	Yes	Yes	Yes
The OMB Generic Method	Yes	Yes	Yes
Extension to the OMB Generic Method	Yes	Yes	No
OMB Questionnaire & Checklist Approach	No	Yes	No
OMB Cost of Quality Failure Approach	No	Yes	No
Thamhain Model	Yes	No	No

The AFSC Method. The AFSC method could lead to continuous process improvement, and the guidance associated with this method seemed to provide sufficient detail to properly exercise the method. Also, the AFSC method seemed to focus on improving individual processes and the method remained a candidate for the development exercise.

Basili & Rombach Paradigm. Although it may very well lead to process improvement, the guidance available on the Basili & Rombach method, especially step two, provided insufficient detail to properly exercise it. This method was removed from further consideration.

Hayes & Miller Process. If exercised properly, this method could lead to continuous process improvement. However, the guidance for the Hayes & Miller method provided insufficient detail to properly exercise it, and it was removed from further consideration. The method did seem likely to improve processes because of the focus on behavior that leads to improvement. However, the requirement of step one to select an acquisition objective and the related behaviors easily encompasses more than a single process.

The Kinlaw Process Improvement Model. The Kinlaw method could lead to continuous process improvement, and the guidance for this method provided sufficient detail to properly exercise it. This method focuses on improving individual processes, and it remained a candidate for the development exercise.

The OMB Generic Method. The OMB Generic Method would lead to continuous process improvement, and the method's guidance provided sufficient detail to exercise it. The generic approach seemed to focus on improving individual processes. The OMB Generic Method remained a candidate for the development exercise.

Extension to the OMB Generic Method. The aggregation of multiple characteristics into a single quality index for the entire process can produce continuous process improvement. However, the aggregation changes the focus from an individual process to the entire system. This may be a good approach for reporting quality to higher

levels of management, but it was removed from further consideration for the metric development exercise.

OMB Questionnaire & Checklist Approach. The guidance for this alternate OMB method provided sufficient detail, but this method would probably not lead to continuous process improvement. Once all of the questions on the checklist were answered affirmatively, this method would no longer produce improvement. It seems best suited to aggregate measures and should not be used for individual metrics. The method was removed from further consideration.

OMB Cost of Quality Failure Approach. This method is not sufficiently focused on the process or the customer. While this metric would help the business bottom line in the short term, it may not drive appropriate behaviors. For example, a change in accounting procedures could actually reduce the cost of quality, but not improve the process. The method was removed from further consideration.

Thamhain Model. The Thamhain model could lead to continuous process improvement, but this method's guidance provided insufficient detail to properly exercise the method. While his questions are invaluable for managers who need to understand their processes, and the resulting answers would provide insight into possible metrics to use, just answering the questions does not represent a step by step methodology that would yield high quality metrics. Also, the model seemed to focus on improving an entire system rather than individual processes and was also removed from further consideration.

Initial Screening Summary. The AFSC method, the Kinlaw process improvement model, and the OMB Generic Method were determined to be the only candidates for the next level of screening.

Secondary Screening

This level of screening was performed on the AFSC method, the Kinlaw process improvement model, and the OMB Generic Method. The methods were rated against the criteria listed in Table 3-2. This screening was done by an expert panel consisting of the research team and the research advisors. Table 4-2 provides a summary of the secondary screening results. Those methods not selected for the metric development exercise are shown as shaded columns in the table.

The AFSC Method. The expert panel rated this method as MODERATELY in six of the ten criteria. The method was determined to be BARELY executable because the guidance does not adequately describe how to generate new metrics. The fourth step of the AFSC method simply states "Generate new metrics if existing metrics are inadequate" (AFSC, 1991:3-3 or 4). *This is a serious omission for a handbook designed* to provide sufficient information to begin developing metrics (AFSC, 1991:1-1). The method was also rated BARELY efficient, primarily because the first step aligns the purpose with the organization's mission, goals, and objectives. This can be so difficult and time consuming that it becomes counterproductive when trying to improve an individual process. Also, the AFSC method was rated as BARELY process-oriented, because it did not require identification of the individual steps and interim outputs of the process producing the output of interest. Finally, the panel rated the method NOT AT ALL sufficient to develop metrics also primarily because no guidance was provided on how to develop metrics in step four. Therefore, the AFSC method was not selected for the development exercise.

Table 4-2

Secondary Screening Results

Criteria	The AFSC Method	The Kinlaw Method	The OMB Generic Method
1. How understandable were the steps? (Was the purpose of each step clear?)	Moderately	Barely	Moderately
2. How executable were the steps? (Could you actually perform them?)	Barely	Moderately	Quite a bit
3. How efficient was the method? (Did each step add value?)	Barely	Moderately	Moderately
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)	Moderately	Moderately	Moderately
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)	Moderately	Moderately	Moderately
6. How sufficient was the method? (Did the guidance list all the steps necessary to develop a metric?)	Not at all	Moderately	Moderately
7. How process-oriented was the method? (Did the guidance force you to focus on a process needing improvement?)	Barely	Barely	Quite a bit
8. How assessable was the result? (Was a mechanism provided to assess the quality of the metric produced?)	Moderately	Barely	Moderately
9. How helpful were the examples provided for each step of the method?	Moderately	Barely	Quite a bit
10. How well did the guidance force you to focus on the customer of the product or service provided?	Moderately	Moderately	Quite a bit

The Kinlaw Process Improvement Model. The panel of experts determined the Kinlaw method MODERATELY satisfied six of the ten criteria, but only barely satisfied four criteria. The method was BARELY understandable due to the confusing structure of the overall method and the wording of the individual steps. The level of detail provided in the method is so great that users would be challenged to understand the

method sufficiently enough to produce high-quality metrics without a process flow diagram. The method was BARELY process-oriented because it did not require a thorough description of the process which produced the outputs of interest. When using the Kinlaw method, care should be taken to identify all internal steps and outputs. The panel rated the method as BARELY assessable because of the lack of criteria with which to assess the metrics produced, and the examples provided were confusing and BARELY helpful. Therefore, the Kinlaw method was not selected for the metric development exercise.

The OMB Generic Method. The OMB Generic Method did not receive any rating less than MODERATELY, and four of the ten criteria were rated as QUITE A BIT. The expert panel determined that the OMB steps were highly executable because of the exacting descriptions and accompanying worksheets. It was also the only method that required a specific identification of all steps in the process of interest and provided a continuous service-oriented example through every step in the method. The method was also rated QUITE A BIT for customer orientation since it was the only one to specifically require identification of the ultimate and all interim customers as well as their requirements and expectations. Therefore, the expert panel selected the OMB Generic Method for the development exercise.

Secondary Screening Summary. The OMB Generic Method was identified as the method most likely to produce high-quality metrics. It was in fact the only method to be rated QUITE A BIT for any criteria and the only method to have no BARELY or NOT AT ALL ratings. The AFSC and Kinlaw methods had no QUITE A BIT ratings and four BARELY or NOT AT ALL ratings. Therefore, the OMB Generic Method was chosen for the metric development exercise.

Metric Development Exercise

The OMB Generic Method was determined to be the method most likely to produce high-quality metrics and was used for the metric development exercise. While the use of this method will not be taught here, a full narrative can be found in chapter two, and a flowchart, as drawn by the researchers, is shown in Figure 4-1. The documentation of all the results of every step in the metric development exercise (which to our knowledge has never before been attempted) provides a record of a complete group process to develop metrics, and helps the readers to understand how to implement metric development in a real-world environment.

Exercise Participants. Five members of the Faculty Qualification and Recruitment Committee at the Air Force Institute of Technology's Graduate School of Logistics and Acquisition Management volunteered to use the OMB Generic Method to develop metrics for the faculty recruitment process. All participants possessed Ph.D. degrees and had completed from zero (2 participants) to more than forty (2 participants) contact hours of education or training on the subject of quality. With the exception of the two research advisors on the committee, no participant was familiar with or had used any metric development methods prior to this exercise. Some participants even indicated no awareness of any of the candidate authors on the subject of quality, while others indicated they were aware of several authors but had actually read works by only a handful. The OMB Generic Method consists of five steps, the results of which will be presented in three phases representing the groupings used in the OMB worksheets and facilitated by the researchers during the exercise.

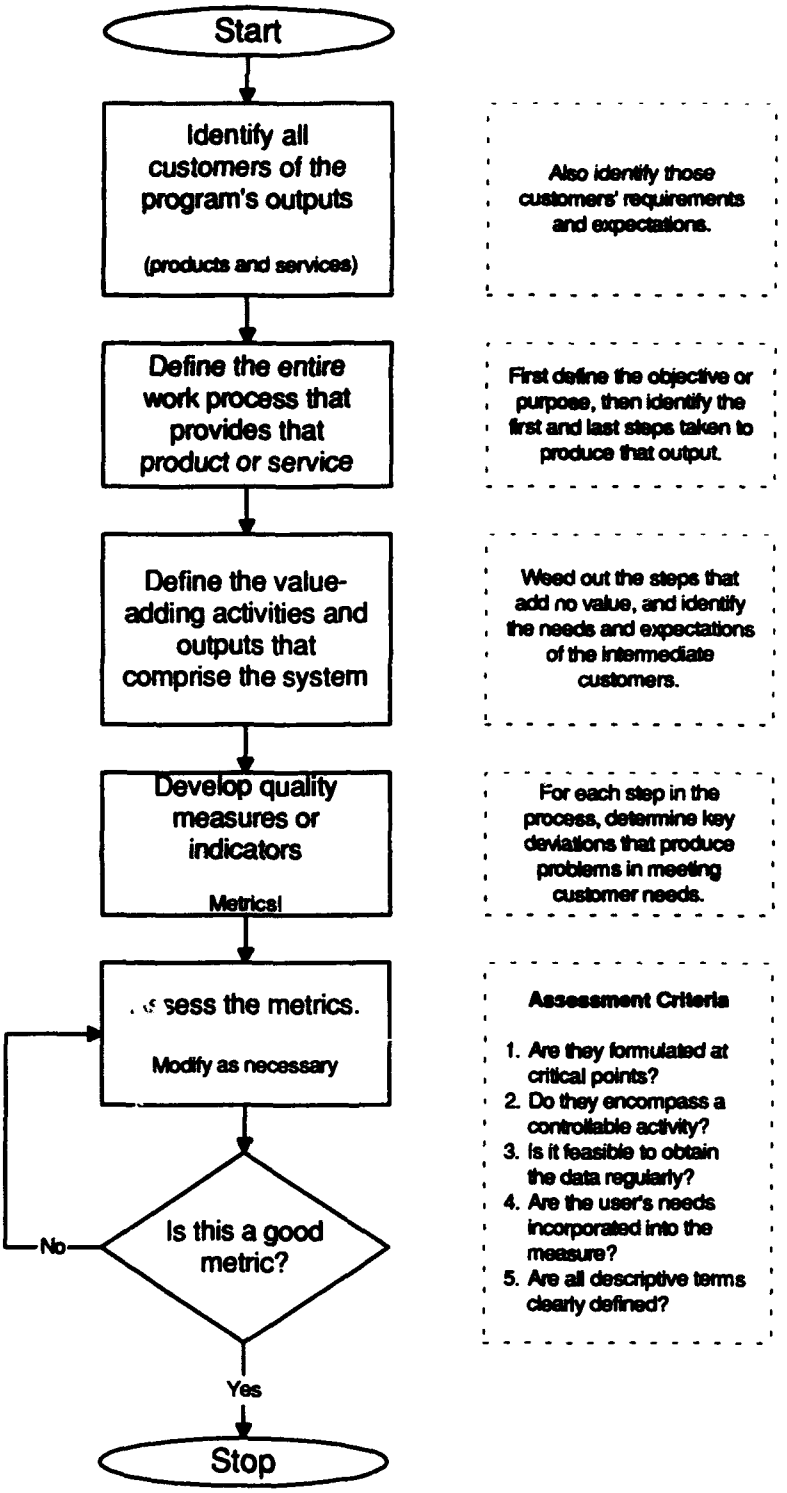


Figure 4-1. The OMB Generic Method

Phase I. Identify the Customers. The OMB Generic Method provided a worksheet to assist in the first step of listing the final products and services produced by a process, and identifying the customers and their requirements and expectations of those products and services (OMB, 1989:30). This act of identifying the outputs, customers and requirements in effect defined the purpose of the process of interest, which was required to begin the next step. The committee identified the output of the faculty recruitment process to be a list of candidates for entry into Doctor of Philosophy (Ph.D.) programs at civilian universities. Then they also determined the customers of that output, and the customers' requirements and expectations as shown in Table 4-3.

Table 4-3

Phase I Results

Final Output	Customers	Requirements and Expectations
Candidate List	Administrators	Retainable, promotable, and available candidates
	Faculty Department heads Dean Students	Relevant experience Academic potential Retainable Promotable Teaching and/or research skills Interpersonal skills Character
	Civilian universities	Academic potential Teaching and/or research skills

Phase II. Define the System. Phase II combined steps two and three of the OMB Generic Method. Step two required participants to define the entire work process by listing the first and last steps in the process (process boundaries) after the objective or purpose of that process was defined. Step three required participants to define all the value-adding activities and outputs of the process, the customers of each activity, and

their requirements and expectations. This step enabled the removal of any activities that did not add value. A worksheet was also provided to assist in listing the steps in the process being improved, and the outputs of those steps (OMB, 1989:31).

The committee first determined the boundaries of the process of interest. They initially determined first and last steps of the process which were too broad in scope, and then limited the boundaries to those activities within their control. Next, the committee identified all value-adding steps between the first and last ones and drew a flowchart of the faculty recruitment process. All the steps of the faculty recruitment process are depicted in Figure 4-2, and the output of each step is listed in Table 4-4. The committee then identified the customers of each output, and the customers' requirements and expectations of those outputs, also shown in Table 4-4.

Phase III. Develop Quality Measures. This phase combines steps four and five of the *OMB Generic Method*. *Step four* required participants to determine the key deviations that produce problems or variations in meeting the customers' needs and expectations, and then to develop metrics for those problem areas. Step five required participants to assess the adequacy of the metrics developed. The metrics were checked against the assessment criteria to ensure they will result in improving the process. A worksheet was provided to assist in listing the key deviations or problems associated with each step identified in Phase II (OMB, 1989:33).

Due to time constraints, the committee limited their efforts to producing metrics for only those steps in the process which had previously been significant sources of

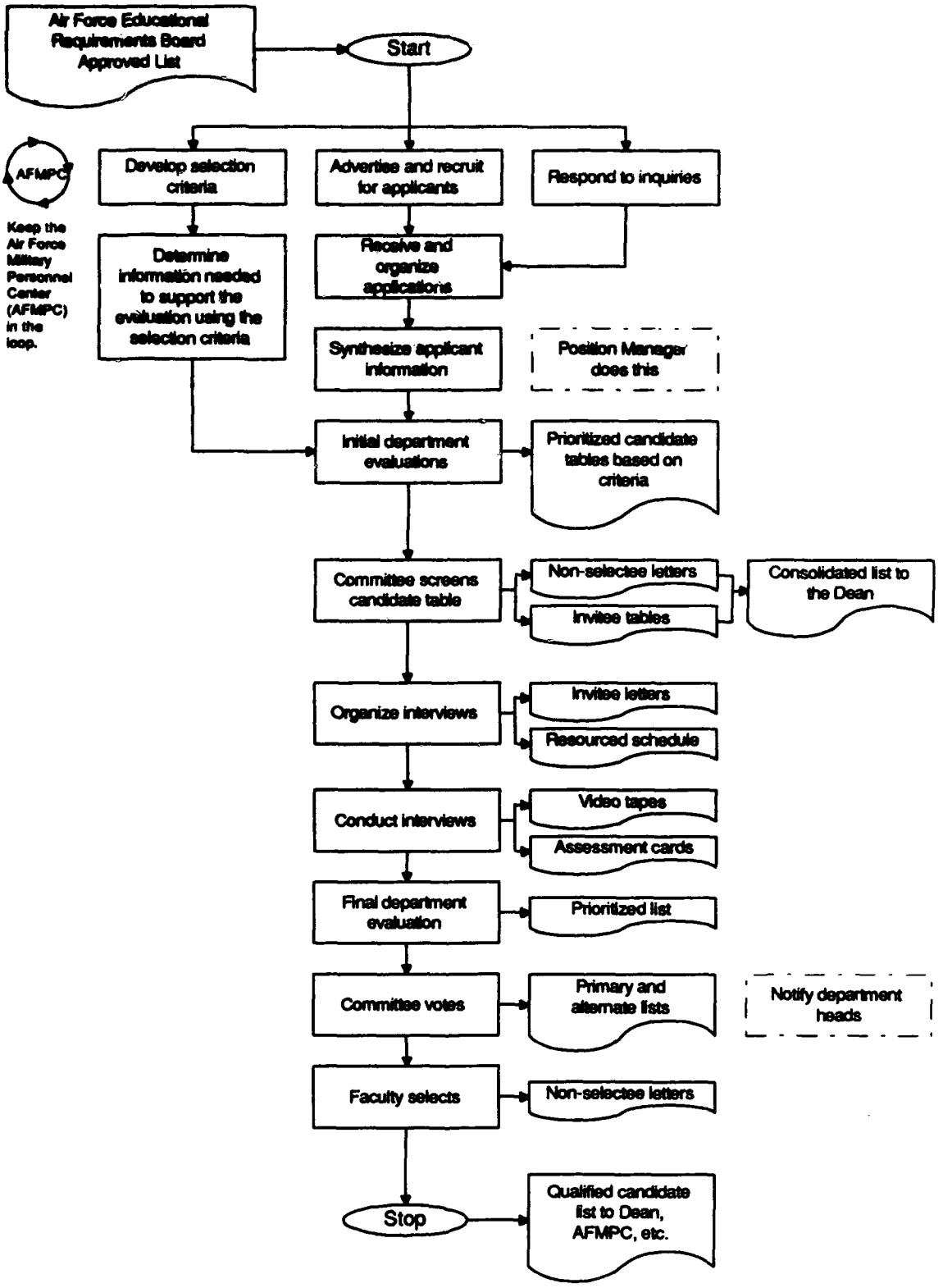


Figure 4-2. The Faculty Recruitment Process

Table 4-4

Phase II Results

Step	Output	Customers	Requirements and Expectations
Develop selection criteria	Selection criteria	Evaluators	Fair, easy, rational, detailed, unbiased, accepted, understandable, practical
Advertise and recruit for applicants	Ads	Potential applicants	Timely information on schools and programs available, the job, how to apply, minimum standards
Respond to inquiries	Answers	Potential applicants	Accurate and timely information
Determine information needed to support the evaluation using the selection criteria	The data required	Evaluators Position managers	Information necessary and sufficient to make a decision
Receive and organize applications	Sorted applicant folders	Position managers	Complete package for all applicants
Synthesize applicant information	Tables	Department heads	Complete and understandable table, reflecting good judgment by the position manager on the subjective criteria
Initial department evaluations	Prioritized list and tables	Committee	Complete data in the proper format which reflects the selection criteria. Department consensus.
Committee screens candidate table	List of invitees	Committee	Complete address/telephone data
	Non-selectee letters	Non-selectees	Courtesy and timeliness
	Consolidated list	Dean	Completeness and timeliness
Organize interviews	Resourced interview schedule. Invitee letters	Faculty, department heads, Dean, invitees	Schedule flexibility, lead time, preparation information
Conduct interviews	Video tape Assessment cards	Committee	Complete, organized, timely, and available video tapes. Legible assessment cards.
Final department evaluation	Prioritized list of candidates	Committee	Complete data in the proper format which reflects the selection criteria. Department consensus.
Committee votes	Identification of primary and alternate choices	Faculty, Department heads	Faculty needs the information necessary to make a decision. Department heads need a qualified candidate.
Faculty selects	Approved candidate list. Non-selectee letters	Administrators, faculty, department heads, Dean, candidates, universities	Qualified candidates

problems and provided the greatest opportunity for process improvement. The deviations and measures identified are listed in Table 4-5.

Table 4-5

Phase III Results

Step	Key Deviations or Problems	Measures	Assessment
Advertise and recruit for applicants	Late ads Incomplete applications, lack of control, unsatisfactory response pool	$\frac{\text{Number of Recruits}}{\text{Total Number of Applicants}}$ $\frac{\text{Number of Qualified Applicants}}{\text{Total Number of Applicants}}$	Both met all criteria
Receive and organize applications	Incomplete packages	$\frac{\text{Number of Complete Packages}}{\text{Total Number of Packages}}$ $\frac{\text{Number of On-time Applications}}{\text{Total Number of Applications}}$	Both met all criteria
Conduct initial department evaluations	Tables still in work, not standard, criteria not applied uniformly	$\frac{\text{Number of Complete Rows}}{\text{Total Number of Rows}}$	Met all criteria
Conduct interviews	Assessment, Low faculty attendance	$\frac{\text{Number of Faculty Viewing Tapes}}{\text{Total Number of Faculty}}$	Met all criteria
Faculty selects	AFMPC rejections	$\frac{\text{Number Accepted by AFMPC}}{\text{Total Number Submitted}}$	Met all criteria
	University rejections	$\frac{\text{Number Accepted by University}}{\text{Total Number Submitted}}$	Met all criteria
	Lack of faculty consensus	$\frac{\text{Number Confirmed by Faculty}}{\text{Number Submitted by Committee}}$ $\frac{\text{Number of Yes Votes}}{\text{Total Number of Votes}}$	Both met all criteria

Analysis of Results. The exercise was highly successful. The participants identified the customers of the faculty recruitment process, defined the entire system, and developed quality measures in only four hours. One factor that reduced the time required to complete the exercise was the high degree of familiarity the committee members demonstrated with the faculty recruitment process. Less familiarity with the process of interest would make the metric development effort more difficult and time consuming. Starting with the identification of the process boundaries, every step in the metric development method fostered a greater understanding of the process of interest.

At one point, the committee realized they had not been systematically recruiting candidates, but had only been advertising for candidates. While developing a measure to reduce the unsatisfactory response pool identified as a problem when advertising for applicants, the committee realized an active approach of recruiting qualified candidates made more sense than the passive approach of advertising for candidates from a pool of both qualified and unqualified individuals. The committee then changed the title of the step from "Advertise" to "Advertise and recruit," which represented a shift toward a new approach which more closely matched the intent of the faculty recruitment process. This resulted in a process improvement without having yet identified a metric.

Measures were also identified which should result in further process improvement. The committee identified as a problem the small number of faculty members reviewing the video tapes of the candidate interviews. The following measure (number of faculty viewing tapes divided by the total number of faculty) was designed to improve that portion of the process, and this measure is an example of a high-quality metric.

$$\frac{\text{Number of Faculty Viewing Tapes}}{\text{Total Number of Faculty}}$$

This is a high-quality metric because it will result in improvement of the process as the ratio gets better (higher or lower depending on the specific measure). The ratio can be improved in one of two ways. Either the denominator (total number of faculty) can be reduced, which may not be desired, or the numerator (number of faculty viewing tapes) can be increased. In this case, increasing the number of faculty members viewing the tapes is best accomplished by increasing the level of faculty interest, which is how this metric leads to process improvement in addition to providing the current status of the process. The metric actually identifies the action required to improve the process — the characteristic that separates a metric from a measurement. The actual implementation of how to increase the level of faculty interest, while important, is external to the actual metric development process, and the committee decided to pursue this effort on their own.

Participant Responses. In the interviews following the metric development exercise, all of the participants stated they were highly (QUITE A BIT) comfortable with using the OMB Generic Method to produce metrics, and four of the five participants stated they were also very (QUITE A BIT) comfortable with the results produced. The fifth participant was MODERATELY comfortable with the results. When asked about strengths of this method, participants identified several strong qualities:

- Quick
- Logical
- Meaningful
- Simple to do
- Ran smoothly
- Process based
- Group approach

- No unnecessary steps
- Forces you to define your process
- Focuses on both external and internal customers
- Yields high-quality metrics at critical points in the process
- Easy to understand by someone unfamiliar with the method

Participants also identified weaknesses for this method and recommended the following improvements:

- Add a formal step for participant preparation (method familiarization and data collection)
- When identifying the customers of the final outputs (step one), prioritize the customers and the outputs.
- When defining the entire work process (step two)
  - Require the process to be drawn as a flowchart.
  - Search for existing models or flowcharts of the process of interest.
  - Emphasize the need to identify process boundaries which yield a controllable work process.
- When developing the metrics (step four) only develop them for the most critical points in the process of interest.
- Emphasize the big picture. Show the metrics in an improvement feedback loop.

The ratings against the secondary screening criteria from Table 3-? assigned by the participants for this method are listed in Table 4-6. The ratings are shown as fractions, where 5/5 indicates five out of five participants assigned a particular rating.

Table 4-6

Participant Assessment of the OMB Generic Method

Criteria	Rating			
	Blue	Green	Yellow	Red
	Quite a bit	Moderately	Barely	Not at all
1. How understandable were the steps? (Was the purpose of each step clear?)	5/5 [4]			
2. How executable were the steps? (Could you actually perform them?)	4/5	1/5		
3. How efficient was the method? (Did each step add value?)	4/5	1/5		
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)	4/5	1/5		
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)	4/5		1/5	
6. How sufficient was the method? (Did the guidance list all the steps necessary to develop a metric?)	4/5	1/5		
7. How process-oriented was the method? (Did the guidance force you to focus on a process needing improvement?)	4/5		1/5	
8. How assessable was the result? (Was a mechanism provided to assess the quality of the metric produced?)	4/5		1/5	
9. How helpful were the examples provided for each step of the method?	3/4 [5]	1/4		
10. How well did the guidance force you to focus on the customer of the product or service provided?	3/5	2/5		

Researchers' Observations. The metric development exercise was a success. Participants were highly satisfied with both the method and the results. The OMB

[4] 5/5 indicates five out of five participants assigned this rating.

[5] One participant did not read the examples provided and therefore chose not to rate this criterion.

Generic Method was demonstrated to be both easy to facilitate and effective in defining all the steps of the faculty recruitment process and allowing participants to develop high-quality metrics which should result in process improvement. Particularly notable was the fact that participants unanimously indicated that the method was very (QUITE A BIT) understandable, even though they had only limited knowledge of this method beforehand. The method does leave some room for improvement. The strengths of this method, and participants' recommendations for improvement, were considered in the development of a proposed metric development method presented in chapter five.

Methods Found in Use at ASC (Garofoli, 1993)

AFMC and ASC use the Corporate Management Network (CMN) to collect metrics of interest to the respective commanders, but many SPO's within ASC have developed internal metrics to manage their own processes because the CMN metrics are not useful to the organizations on a daily basis. In fact, one SPO reported that they "only use one of the twenty-five" CMN metrics internally. Captain Dave Garofoli, the Director of Measurement at the ASC Total Quality Office (ASC/TQ), identified organizations within ASC that had undergone metric development exercises within the last twelve months. Three organizations agreed to participate in this field study, each of which used a different metric development method.

The AFSC Method (Dierker, 1993). The AFSC Method was used to develop metrics common to functional areas within ASC. Those functional areas were the Directorates of Acquisition Logistics (ASC/AL), Contracting (ASC/PK), Engineering (ASC/EN), Financial Management (ASC/FM), and Program Management (ASC/CY). Refer to chapter two for a complete description of the AFSC method. The metric development effort was facilitated by Mr. Greg Dierker, an operations research analyst

within ASC/AL, who indicated the metrics developed were "fairly new" metrics representing a "hybrid" of the metrics currently used by the five functional areas. The processes of interest common to the five functional areas were the management of (1) meeting the certification requirements of the Acquisition Professional Development Program (APDP), (2) meeting the ASC policy for Company Grade Officer (CGO) professional development by rotating those officers throughout different jobs within ASC, (3) the use of Scientific and Engineering Technical Assistance (SETA) contracts within ASC, and (4) meeting the ASC goal of giving promotion-eligible lieutenant colonels the best opportunity for advancement by ensuring they are in promotable positions. The first five steps of the AFSC method encompass the development of a metric, while the remaining five steps are concerned with the presentation and use of the metric (AFSC, 1991:3-3 to 3-5). Figure 4-3 is a flowchart of the first five steps of the AFSC Method as drawn by the researchers.

Participant Experience Levels. Four persons who had used the AFSC method were interviewed. All of them had taken undergraduate courses, and one individual had received an undergraduate degree. Three of the four had more than forty contact hours of education or training on the subject of quality, while the fourth participant had between eight and forty hours. No participant was familiar with or had used any metric development method other than the AFSC method. Some participants indicated no awareness of any of the candidate authors on the subject of quality, while others indicated they were aware of a few authors but had actually read works by at most two of them.

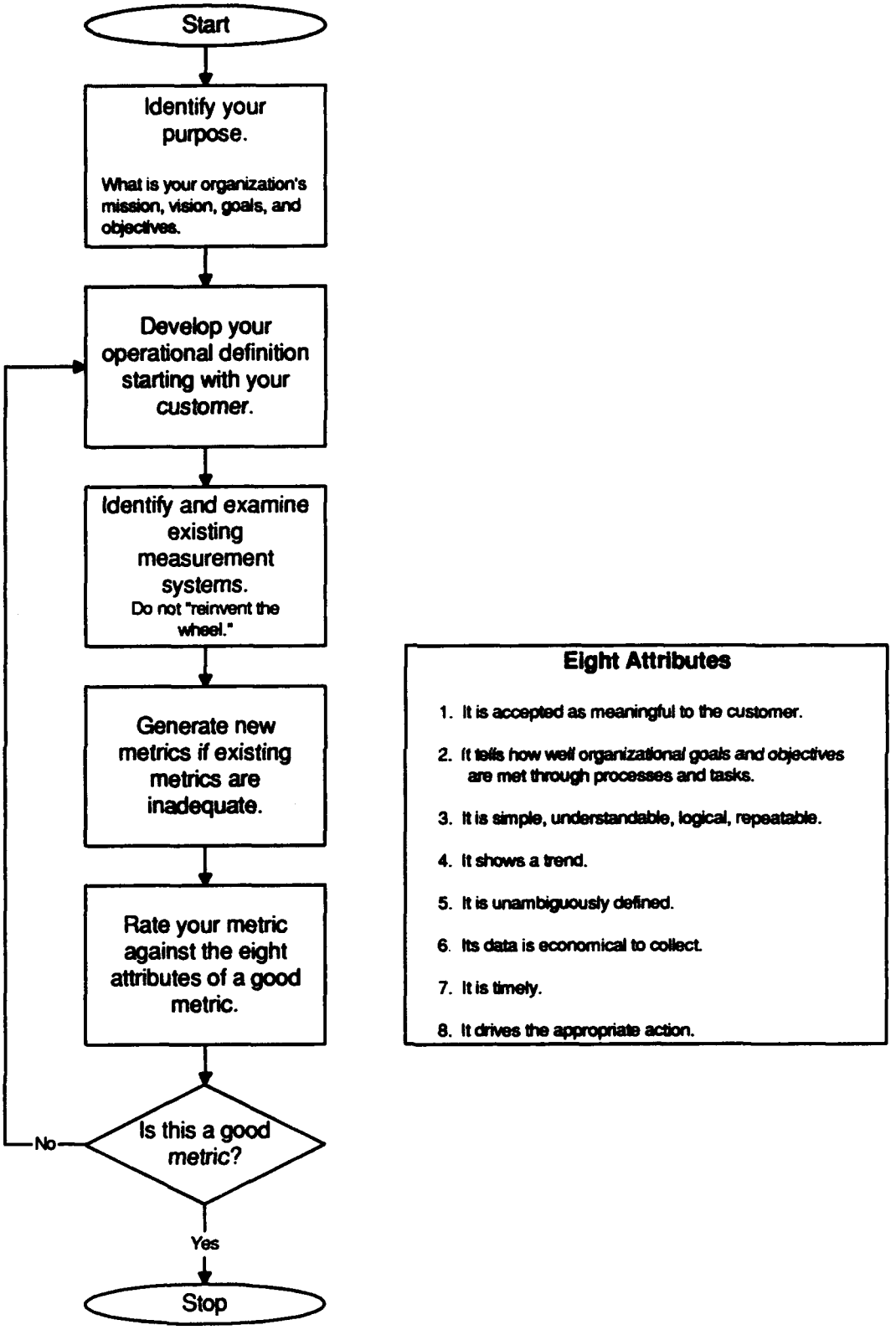


Figure 4-3. The AFSC Method

Participant Responses. Three of the participants stated they were highly (QUITE A BIT) comfortable with using the AFSC method to produce metrics, and three of the four participants stated they were very (QUITE A BIT) comfortable with the results produced. One participant was MODERATELY comfortable with using the AFSC method to produce metrics, and one participant was MODERATELY comfortable with the results produced. When asked about strengths of this method, participants identified:

- Common sense steps
- Great assessment step
- Method gives something to go by
- Forces you to look at the purpose
- Saves time by looking at existing metrics
- Forces you to look at process, customers, requirements, and improvements

Participants also identified weaknesses for this method and recommended the following improvements:

- Get the process owner involved
- Use more examples for the steps
- Get agreement on the purpose of the process or function of interest
- Have management support and empower the metric development team
- Make the steps easier to understand, especially the operational definition step
- Add a step to validate that the metric is still useful after improvement efforts are in place

The ratings assigned by the participants for this method are listed in Table 4-7. The ratings are shown as fractions, where 3/4 indicates three out of four participants assigned a particular rating.

Researchers' Observations. Participants were generally satisfied with both the method and the results. Empirically, these results appear stronger than those for the OMB Generic Method. However, recall that participants were only familiar with this one metric development method. In addition, their metrics tended to be status reporting tools which would not necessarily lead to continuous improvement because the metrics were goal-oriented rather than process-oriented. The goal was usually 100 percent compliance with a simple policy, and the metric merely reported progress toward that goal. Unfortunately, the metrics gave no insight into how the process itself could be improved. In fact, these metrics are fairly typical of many ASC organizations who are using high-powered measurement tools to track very simple problems.

For example, all the CGO Rotation metric shows is that CGOs are either moved to a new job within three years or not. The solution to less than 100 percent performance is then to move more CGOs. While the AFSC Method might allow participants to develop metrics which could result in process improvement, the method has some major drawbacks. The strengths of this method, and participants' recommendations for improvement, were considered in the development of a proposed metric development method presented in chapter five.

Table 4-7

Participant Assessment of the AFSC Method

Criteria	Rating			
	Blue	Green	Yellow	Red
	Quite a bit	Moderately	Barely	Not at all
1. How understandable were the steps? (Was the purpose of each step clear?)	3/4 [6]	1/4		
2. How executable were the steps? (Could you actually perform them?)	2/4	2/4		
3. How efficient was the method? (Did each step add value?)	2/4	2/4		
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)	3/4	1/4		
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)	3/4	1/4		
6. How sufficient was the method? (Did the guidance list all the steps necessary to develop a metric?)	3/4	1/4		
7. How process-oriented was the method? (Did the guidance force you to focus on a process needing improvement?)	2/4	2/4		
8. How assessable was the result? (Was a mechanism provided to assess the quality of the metric produced?)	3/4	1/4		
9. How helpful were the examples provided for each step of the method?	2/3 [7]	1/3		
10. How well did the guidance force you to focus on the customer of the product or service provided?	3/4	1/4		

[6] 3/4 indicates three out of four participants assigned this rating.

[7] One participant did not read the examples provided and therefore chose not to rate this criterion.

The ASC/SD Method (Buchanan, 1993; Conley, 1993). The Aircraft Systems Program Office (ASC/SD) within ASC implemented a different approach to developing metrics. This organization is a basket SPO, housing program offices for twenty-four different types of aircraft. The SPO director sought to standardize the presentation of metrics common to the twenty-four programs, and established a Tiger Team to perform that standardization, initially for the areas of cost, schedule, performance, and supportability.

The SPO director identified the objective for each type (area) of metric, and the Tiger Team divided up into smaller teams to brainstorm new metrics for that area, or to synthesize metrics already in existence within the twenty-four programs into a common metric for that area. The Tiger Team, SPO Director, and the Integrated Product Team (IPT) Leader for each of the twenty-four programs would then review, coordinate, and modify the common metric as necessary. Figure 4-4 is a flowchart of the ASC/SD Method as drawn by the researchers.

Participant Experience Levels. Three persons who had used the ASC/SD method were interviewed. Two of them had received undergraduate degrees, and one individual had received a graduate degree. One of the participants had more than forty contact hours of education or training on the subject of quality, while the others had between eight and forty hours. No participant was familiar with or had used any metric development method other than the ASC/SD method. All participants indicated awareness of at least one of the candidate authors on the subject of quality, and one participant had actually read works by four of the authors.

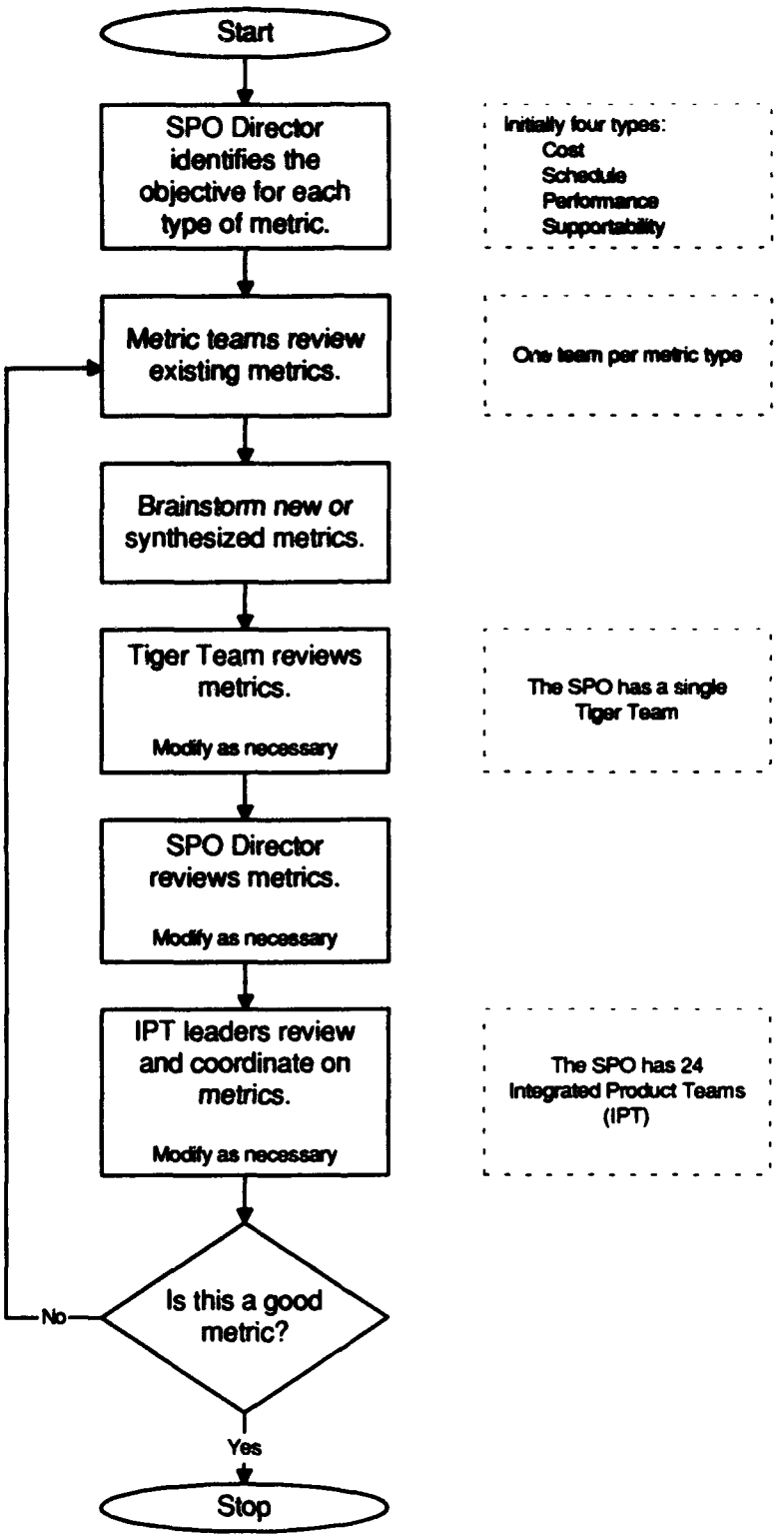


Figure 4-4. The ASC/SD Method

Participant Responses. Two of the participants stated they were MODERATELY comfortable with using the ASC/SD method to produce metrics, and the other participant was BARELY comfortable with using the method. All three participants were MODERATELY comfortable with the results produced. When asked about strengths of this method, participants mostly identified the following qualities:

- Review process
- Brainstorming sessions
- Identification of the objective
- Strong involvement from SPO Director
- Proper execution will lead to good measures
- Team gets better understanding of how other members operate in the SPO

Participants also identified weaknesses for this method and recommended the following improvements:

- Get sharp people
- Know what the real objectives are
- Identify the mission, vision, and goal of the organization
- Use formal team-building for the metric development team
- Have a realistic schedule for completing the review process
- Have management support and empower the metric development team

The ratings assigned by the participants for this method are listed in Table 4-8. The ratings are shown as fractions, where 3/3 indicates three out of three participants assigned a particular rating.

Table 4-8

Participant Assessment of the ASC/SD Method

Criteria	Rating			
	Blue	Green	Yellow	Red
	Quite a bit	Moderately	Barely	Not at all
1. How understandable were the steps? (Was the purpose of each step clear?)	3/3 [8]			
2. How executable were the steps? (Could you actually perform them?)	1/3	2/3		
3. How efficient was the method? (Did each step add value?)		2/3	1/3	
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)	3/3			
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)	2/3	1/3		
6. How sufficient was the method? (Did the guidance list all the steps necessary to develop a metric?)	1/3	2/3		
7. How process-oriented was the method? (Did the guidance force you to focus on a process needing improvement?)	2/3	1/3		
8. How assessable was the result? (Was a mechanism provided to assess the quality of the metric produced?)		3/3		
9. How helpful were the examples provided for each step of the method?	[9]			
10. How well did the guidance force you to focus on the customer of the product or service provided?	2/3	1/3		

Researchers' Observations. Participants were generally satisfied with both the method and the results. However, this metric development effort was plagued by

[8] 3/3 indicates three out of three participants assigned this rating.
[9] No examples were provided for this method.

several notable problems. First, it seemed to concentrate more on standardizing the presentation of current metrics rather than the development of new metrics that would foster process improvement. Unfortunately, the current metrics also tended to be status reporting tools which were goal-oriented rather than process-oriented. Again, the metrics merely reported progress toward a particular goal and gave no particular insight into how the process itself could be improved.

Another problem was the difficulty of getting volunteers to participate in the Tiger Team in any capacity. When asked about this, one member of the Tiger Team stated there was "a lot of inertia to metrics" within the organization (Buchanan, 1993). This resulted in most of the work being done by a small cadre without representation from each IPT or support from the IPT leaders. Therefore the coordination and review process was time consuming and conflict-filled. In fact, the organization has only completed two of the four types of metrics they set out to standardize in June of 1992.

The ASC/SD Method is not likely to help participants develop metrics which result in process improvement. Of course, competent participants may arrive at high-quality metrics in spite of the method. The strengths of this method, and participants' recommendations for improvement, were considered in the development of a proposed metric development method presented in chapter five.

The ASC/VC Method (Pytlik, 1993). The Advanced Cruise Missile (ACM) SPO (ASC/VC) within ASC had developed a stand-alone Acquisition Program Tracking System (APTS) which, as a by-product, incorporated metrics (ACM, 1993:3). The System Program Director (SPD) stated that one "shouldn't go in to develop metrics" and that "if the metrics were not used day-to-day, then they were not worthwhile." The director also informed his people that "looking good is not the key — be honest."

The SPO consisted of several Integrated Product Teams (IPTs) for various functions, including depot activation, program close-out, software deliveries, and weapon

system support. The method required the IPT to outline the scope of its effort, to develop a charter which would allow the team to accomplish that effort, and then to develop the objectives necessary to live up to the team's charter. The team then developed a detailed plan to meet those objectives. This integrated plan included identification of any risks, any assumptions made, management reserves, and the thresholds (trip wires) necessitating team leader attention and/or program director attention. The teams then asked themselves how to determine if they were meeting their objectives, and derived the metric(s) used to manage the process of interest by analyzing the answer(s).

One unique feature of this method was the level of automation used to manage the functions of interest. Commercial off-the-shelf software products running under the *Microsoft® Windows™* operating system[7] on a local area network allowed day-to-day management within the IPT, and also provided the SPD the opportunity to receive updated information on a timely basis. Figure 4-5 is a flowchart of the ASC/VC Method as drawn by the researchers.

Participant Experience Levels. Four persons who had used the ASC/VC method were interviewed. Three of them had received graduate degrees, and one individual had received an undergraduate degree. Two of the four had more than forty contact hours of education or training on the subject of quality, while the other participants had between eight and forty hours. Two participants were familiar with the ASC guidance on metrics, and one had used the ASC guidance to develop metrics. The other participants were not familiar with, nor had they used, any metric development method other than the ASC/VC method. Participants indicated awareness of between

[7]*Microsoft* is a registered trademark and *Windows* is a trademark of Microsoft Corporation.

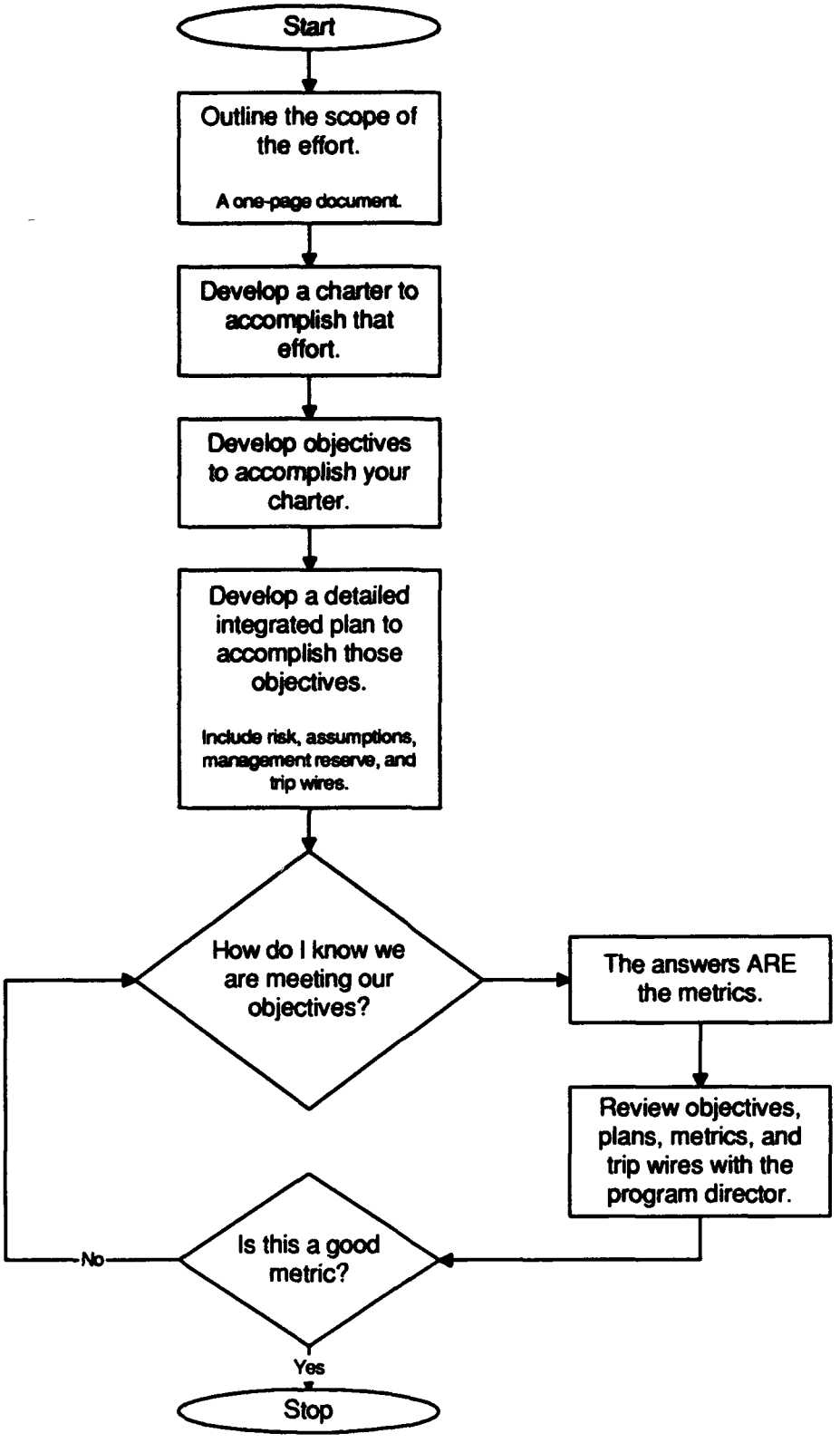


Figure 4-5. The ASC/VC Method

one and four of the candidate authors on the subject of quality, and one participant had actually read works by two of those authors.

Participant Responses. Two of the participants stated they were very (QUITE A BIT) comfortable with using the ASC/VC method to produce metrics, and the other participants stated they were MODERATELY comfortable. Three participants were MODERATELY comfortable with the results produced, and the other participant was very (QUITE A BIT) comfortable with the results. When asked about strengths of this method, participants identified the following:

- It works
- Logical flow
- Top-down approach
- Aimed at meeting objectives
- Automation makes updates easy
- Forces identification of management reserve
- Forces planning, and measurement against that plan
- Well-defined metrics allow keeping a handle on the program
- Standard look of metrics helps management see and understand

Respondents also commented favorably on the usefulness of the ACM APTS automation to manage their functions, and the willingness of the SPD to fund the necessary investments in computer hardware, software, and training.

Participants also identified weaknesses for this method and recommended the following improvements:

- Provide training on metrics before developing them
- Implement metrics at multiple levels of management

- Focus more on the workers' needs rather than the managers'
- In addition to the team leader, allow team members to participate in defining the scope, objectives, etc.
- Add a step to define the goals after the charter is developed and before the objectives are developed

The ratings assigned by the participants for this method are listed in Table 4-9. The ratings are shown as fractions, where 3/4 indicates three out of four participants assigned a particular rating.

Researchers' Observations. Participants were generally satisfied with both the method and the results. The in-house ASC/VC Method does allow participants to develop metrics which should result in process improvement. The key metric development question of "How do I know if I am meeting the objective?" was especially insightful. Ideally, that question should be asked at every value-adding step in the process of interest in order to truly foster continuous improvement. A drawback to this method is the metrics are only as good as the defined objectives. If the objectives are not stated in terms of the customer and are not focused on all the intermediate steps or objectives of a given process, the resulting metrics would not necessarily result in continuous improvement. Also, the determination of the adequacy of the metric developed is only based on discussions with the SPD, rather than formal assessment criteria.

The strengths of this method, and participants' recommendations for improvement, were considered in the development of a proposed metric development method presented in chapter five. While the ACM SPO effectively implemented the metrics into their management system, implementation issues were not the focus of this

Table 4-9

Participant Assessment of the ASC/VC Method

Criteria	Rating			
	Blue	Green	Yellow	Red
	Quite a bit	Moderately	Barely	Not at all
1. How understandable were the steps? (Was the purpose of each step clear?)	3/4 [11]	1/4		
2. How executable were the steps? (Could you actually perform them?)	3/4		1/4	
3. How efficient was the method? (Did each step add value?)	3/4	1/4		
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)	2/4	2/4		
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)	1/4	2/4	1/4	
6. How sufficient was the method? (Did the guidance list all the steps necessary to develop a metric?)		3/4	1/4	
7. How process-oriented was the method? (Did the guidance force you to focus on a process needing improvement?)	1/4	2/4	1/4	
8. How assessable was the result? (Was a mechanism provided to assess the quality of the metric produced?)		2/4	2/4	
9. How helpful were the examples provided for each step of the method?		4/4		
10. How well did the guidance force you to focus on the customer of the product or service provided?	2/4		2/4	

research. The ACM SPO would serve well as a case study in the use of metrics in day-to-day management.

[11] 3/4 indicates three out of four participants assigned this rating.

Metric Awareness

This research effort identified a lack of understanding about metrics and metric development on the part of the participants in this effort. For example, when asked to define the term metric, as it relates to quality, more than half of the seventeen respondents failed to mention "process improvement" as part of their definition. Of these responses, the phrase "progress toward a goal" illustrates the common misconception. This misconception might explain why most of the internal SPO metrics explored in this research effort were status-oriented and tracked "progress toward a goal." While progress toward a goal is a worthwhile activity, it is a necessary, but not sufficient condition for a metric — a metric is a means to achieving objectives and goals through process improvement (AFSC, 1991:2-2). Recall that a metric is a "measurement made over time, which communicates vital information about the quality of a process, activity, or resource" (AFSC, 1991:2-1). Metrics should specifically facilitate quality improvement, by giving understanding to processes and their capabilities (AFSC, 1991:1-2).

Also, while the researchers found several methods in the literature that were intended to help organizations develop metrics, survey responses indicated a general lack of awareness of any metric development methods. Four of the sixteen participants in actual metric development efforts were only familiar with one method, and ten participants had no knowledge of any method whatsoever. Efforts to educate the workforce on the existence and purpose of metric development methods would increase the likelihood of successful metric development efforts that result in continuous process improvement.

Summary

Canvassing both the literature and field organizations as sources of metric development methods provided an opportunity to uncover as many methods as possible. After developing ten criteria for assessing a metric development method, the researchers determined that, of the methods documented in the literature, the OMB Generic Method was the most likely to produce high-quality metrics. The researchers also demonstrated that the OMB Generic Method worked very well in practice when they facilitated the use of that method to develop actual metrics for AFIT's Graduate School of Logistics and Acquisition Management Faculty Recruitment Committee.

The researchers analyzed participants' indications of satisfaction, identification of strengths and weaknesses for the metric development methods, and recommendations for improvement to the methods. Based on this analysis, the researchers concluded that none of the other metric development methods found in use at ASC appeared to be as likely to produce high-quality metrics as the OMB Generic Method. The participants in the AFSC and ASC/SD methods were generally satisfied with the methods and the results produced, but this may be because they were either synthesizing existing metrics or trying to measure progress toward a goal for relatively simple processes. The ASC/VC method demonstrated more potential to produce high-quality metrics as long as the participants focus on all the intermediate steps or objectives of a given process. Participants' satisfaction and recommendations for improvement with each method were considered in identifying the key components of metric development, resulting in the metric development method proposed in chapter five, which should produce higher quality metrics than any of the methods currently used.

V. Conclusions & Recommendations

Introduction

This chapter presents the researchers' conclusions on metrics development and their recommendations for nonmanufacturing organizations and individuals interested in metrics development. The conclusions offer answers to the original research questions presented in chapter one that guided the research effort. These conclusions specifically address the benefits of a methodical approach to developing metrics; the value of the OMB Generic Method, which was selected as most likely to yield high-quality metrics; and a review of the desirable attributes of any metric development method. The recommendations include current applications for this research and ideas for future research in the metrics development area.

Conclusions

Measurement. The research team began by investigating the first two research questions: why measurement is important to process improvement and what features of the process should be measured. The team found that measurement is the process of examining certain characteristics of interest. Also, any continuous improvement effort must start with measurement, which is so fundamental to this process that:

> If you cannot measure it, you cannot control it. If you cannot control it, you cannot manage it — it's as simple as that. Measurement truly separates a successful improvement process from one that fails. (Talley, 1991:xi)

One way to prevent process improvement is to measure the wrong features (Tenner and DeToro, 1992:107). This draws management attention away from what should be measured, and wastes valuable resources collecting measurement data that will not lead

to improvement of the process. What should be measured? Simply put, "measure what's important" (Peters, 1987:483). For example, if customers are important to an organization, which they always are in a service industry, then the organization should measure features of interest to the customers.

Measurement and Metrics. The research team then investigated the next two research questions to find the distinction between a measure and a metric and to identify the attributes of a good metric. A measurement may not always communicate vital information about a process or lead to an improved process, but a metric always will. A metric is defined as "a measurement made over time, which communicates vital information about the quality of a process, activity, or resource." A metric is not an end to itself, but rather a means to achieving objectives and goals through quality improvement by giving understanding to processes and their capabilities (AFSC, 1991:1-1 through 2-2). Unfortunately, all the metrics in use today do not necessarily lead to continuous improvement, but good ones will. The researchers identified the following attributes of a good metric (AFSC, 1991:2-1; OMB, 1989:16).

- It is timely
- It drives the appropriate behavior
- It is repeatable and shows a trend
- It encompasses a controllable activity
- It is feasible to obtain the data regularly
- It is accepted as meaningful to the customer
- It is formulated at a critical point in the process
- It is simple, understandable, and unambiguously defined
- It tells how well organizational goals and objectives are met through processes and tasks

Metric Development Methods. The research team's next two research questions sought metric development methods documented in the literature or that had been used at ASC within the past twelve months. The following methods were found in the literature:

- The AFSC Method
- Basili & Rombach Paradigm
- Hayes & Miller Process
- The Kinlaw Process Improvement Model
- The OMB Generic Method
- Extension to the OMB Generic Method
- OMB Questionnaire & Checklist Approach
- OMB Cost of Quality Failure Approach
- Thamhain Model

The following methods had been used at ASC within the last twelve months.

- The AFSC Method
- The ASC/SD Method
- The ASC/VC Method

Metrics Development. As the team answered research question seven and began to identify criteria for a good development method, it quickly became apparent that measures of quality are not easy to develop in service organizations and that ASC in particular has had difficulty developing metrics that truly lead to continuous process improvement. The general consensus was that the majority of the metrics data that is collected in ASC is done not for the benefit of the SPO or functional area, but to send as program status to higher levels of command (Garofoli, 1993). Many organizations have created their own internal metrics with which to manage their processes, but many of the

metrics are not necessarily driving the types of behavior that would lead to continuous improvement. The internal metrics tended to resemble the CMN and command metrics in that they were output or program status-oriented rather than process-oriented. Since metrics evaluation research had been accomplished previously, and a separate evaluation effort was currently underway, this research team explored the possibility of finding a better process with which to develop metrics in the first place.

Evidence from the literature search and interviews of participants in their respective metric development activities strongly suggests that a methodical or systematic approach should be used to arrive at an organization's metrics. A methodical approach specifically guides participants through logical steps that help them arrive at high-quality measures, and results in increased satisfaction of the metric development process participants. The common approaches of borrowing someone else's metric, making something up, or using whatever is currently measured can certainly be successful, but they are analogous to rolling dice and shooting in the dark. Obviously, the better approach is to use a method that is highly likely to result in high-quality metrics every time for a wide variety of processes or functions of interest.

Another research objective was to identify criteria for a good metric development method. Even though several metric development methods were found in the literature, there was no comparison of the effectiveness of those methods or any criteria to use in judging the value of existing methods or creating a new method. Therefore, the researchers developed the following criteria to help ensure that a development method will yield high-quality metrics.

- Assessable Results. The method should provide some way to assess the value of the metrics that are produced.

- Customer-Oriented. The method should encourage the developers to focus on the customers who are both internal and external to the process in question.

- Efficient. Each step should add value or provide a clear benefit to the metric development process without being redundant or creating intermediate products that are not used.

- Executable. Each step should reflect a task that can actually be accomplished within reasonable constraints (time, data, skill levels).

- Flexible. The method should be flexible enough to be used in different settings. Some developers may be able to exercise the method at one sitting for example. Others may need to split the task into separate sessions.

- Helpful Tools. Any tools (examples, worksheets, etc.) provided to aid the developers in performing each step should be helpful.

- Logical. The method should have a logical flow from one step to the next with the output of one step contributing to the execution of the next one.

- Process-Oriented. The method should encourage the developers to fully define the process rather than just focusing on the final results or outputs. The outputs can only be improved by improving the process.

- Sufficient. The method should include everything necessary to yield high-quality metrics and not omit key steps, instructions, definitions, or examples.

- Understandable. Each step in the method should be clear enough to understand exactly what should be accomplished in that step and why it should be accomplished.

These criteria were the product of months of intense study of 1) the principles of quality, 2) the concept of continuous improvement, and 3) metrics and metric development methods in general.

The OMB Generic Method. The researchers' final objective was to identify which of the existing metric development methods would most likely produce high-quality metrics. To accomplish this, they convened an expert panel that used the above criteria to evaluate the methods which survived the initial screening criteria. The panel found that the OMB Generic Method was the most likely to produce high-quality metrics. The researchers then facilitated the use of this method in a group metric development exercise to provide hands-on experience with the method and produce actual metrics for the AFIT Graduate School of Logistics and Acquisition Management Faculty Recruitment Committee.

The metric development exercise was a success. Participants were highly satisfied with both the method and the results. The OMB Generic Method was demonstrated to be very easy to facilitate and very effective in allowing participants to develop high-quality metrics which should result in process improvement. This exercise allowed the researchers to confirm the finding of the expert panel that the OMB Generic Method was highly likely to produce high-quality metrics and provided valuable insight into the metric development process. The OMB Generic Method would therefore be an excellent method to be used by nonmanufacturing organizations desiring to produce metrics that will lead to continuous improvement.

Proposed Metric Development Method. The researchers considered the metric development exercise and the participants' satisfaction with and recommendations for improvement to their respective methods when identifying the necessary steps of a sound metric development method. The OMB Generic Method contains most of them, but the following improvements would result in a metric development method that would be even more likely to produce high-quality metrics than any of the methods described in this research.

- Education. Before any metric development effort occurs, ensure all participants have been educated on metrics in general, and on the characteristics of a good metric in particular.

- Process boundaries. When identifying the customers of a program's outputs (OMB Step One), ensure those outputs are from a process which encompasses a controllable function or work activity. This is one of the assessment criteria listed in the OMB Generic Method, and informing the developer of this requirement up-front should increase the effectiveness of the method (OMB, 1989:16).

- Prioritize customers. After identifying the customers and their requirements and expectations of a program's outputs (OMB Step One), prioritizing the customers may identify the most important requirements and expectations. Focusing on developing metrics for those high-priority ones first should increase the efficiency of the method.

- Draw a flowchart. When defining the entire work process that provides a product or service (OMB Step Two) the process should be drawn graphically to reduce the risk of defining the process inaccurately.

- Identify the greatest opportunities for process improvement. When developing quality measures or indicators (OMB Step Four), first produce metrics for those steps in the process which provide the greatest opportunity for process improvement (i.e. known sources of problems). Focusing first on developing metrics for those activities should increase the efficiency of the method.

- Improve the assessment criteria. Include all of the following attributes of a good metric when assessing the metrics developed (AFSC, 1991:2-1; OMB, 1989:16).

    - It is timely
    - It drives the appropriate behavior
    - It is repeatable and shows a trend
    - It encompasses a controllable activity
    - It is feasible to obtain the data regularly
    - It is accepted as meaningful to the customer
    - It is formulated at a critical point in the process
    - It is simple, understandable, and unambiguously defined

- Implement and validate the metrics. Emphasize the use of the metrics in a process improvement framework. Metrics also need to be reviewed periodically to validate their usefulness in driving process improvement.

A flowchart of the proposed metric development method, containing the recommended improvements to the OMB Generic Method, is shown as Figure 5-1.

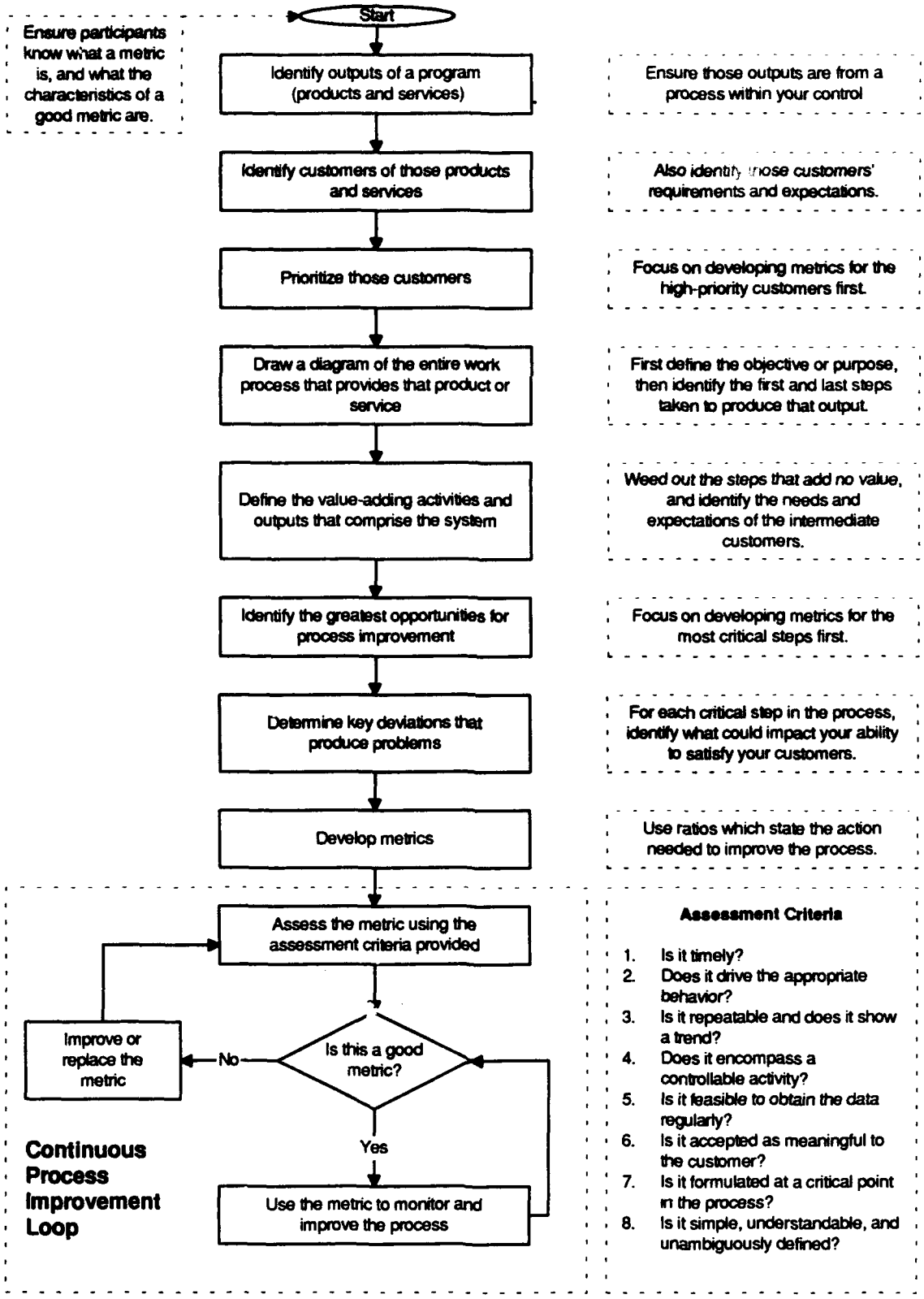


Figure 5-1. The Proposed Metric Development Method

Recommendations for Metric Development

The results of this research can have a valuable impact on metrics development efforts in nonmanufacturing organizations. The researchers recommend the following:

- Any organization wanting to develop internal metrics should begin with either the OMB Generic Method or the alternative one proposed above.

- Quality training functions should first educate the workforce on the true definition and purpose of a metric. Then, they should increase the awareness of the better metric development methods and either teach the workforce how to use the methods or train a core group of people who would facilitate metric development in their organizations. Of the seventeen people interviewed, very few participants in actual metric development efforts were familiar with one method — the AFSC — and most participants had no knowledge of any method whatsoever. Efforts to educate the workforce on the existence and purpose of metric development methods would increase the likelihood of successful metric development efforts that result in continuous process improvement.

- AFMC should replace or modify the current metric development method described in its Metrics Handbook with the OMB Generic Method or the proposed method described above. Either one is likely to yield much better results, both in terms of the quality of the metrics and the satisfaction of the people involved.

- AFMC and ASC should investigate reducing the number and improving the quality of the metrics tracked in their respective systems. There are too many measures disguised as metrics that do not add value to managing the

work performed at ASC. In this era of rising costs, lower budgets, and dwindling manning, no organization can afford the luxury of tracking measures from which no benefit is derived. For example, one organization reported that half of the metrics workload supported requests from ASC and AFMC; only half supported local process improvement. Also, award winning companies like Texas Instruments and Federal Express have shown that fewer, not more, measures should be used as information flows to the upper levels of management (Hummel, 1993; Evans and Lindsey, 1993:131). The extension to the OMB Generic Method may fulfill this requirement in AFMC and ASC as well.

Recommendations for Future Research

- Validate the proposed changes to the OMB Generic Method. New research could incorporate larger sample sizes, metrics for different processes of interest, and evaluations of the metrics produced.

- Perform a comparative analysis experiment pitting two methods against each other to see which yields the best metrics.

- Perform a feasibility study on which steps of the proposed metric development process can be automated. What steps can be performed using a Group Decision Support System (GDSS) like the one available at Armstrong Laboratory at Wright-Patterson AFB?

- Study the CMN system and the metrics required by AFMC. Which metrics only report program status? Which ones can be deleted? Which ones truly

drive performance that leads to continuous improvement? What information does AFMC really need? How can it all be improved?

- The ACM SPO effectively integrated the use of metrics into their management system as a by-product of their improvement process. However, the use of metrics to manage and improve a process was not the focus of this research. The ACM SPO would serve well as a case study in the use of metrics in day-to-day management.

- Investigate the impression of metrics in AFMC, including definitions, usefulness, and suggestions for improvement.

Summary

The purpose of this research was to explore the use of metric development methods in nonmanufacturing organizations. Several methods were found in the literature and in use within ASC. The OMB Generic Method was identified as most likely to result in high-quality metrics according to the secondary screening criteria developed by the researchers and used in an actual metrics development exercise. Interviews with participants of that exercise and the metric development efforts in ASC were conducted and considered in identifying the necessary steps of a sound metric development method. Those steps were incorporated into a proposed metric development method that was based on the OMB Generic Method, and should be more likely to produce high-quality metrics that will result in continuous process improvement. The researchers also identified problems in AFMC metrics use and recommended corrective actions and ideas for future research.

Appendix A. Survey Of Experience

Survey of Experience

The following information will be used to report individual skill levels, experience levels, and other demographic characteristics of the individuals participating in this research effort.

Name:_____ Age:_____ Rank:_____

Present Job Title:_____ Years in present job:_____

1. What is the highest level of education you have attained? (Mark one)

   - ☐ Some High School
   - ☐ Completed High School
   - ☐ Some Undergraduate College Courses
   - ☐ Undergraduate Degree
   - ☐ Some Graduate-Level Courses
   - ☐ Graduate Degree
   - ☐ Doctoral Degree

2. How much education and/or training have you received on the subject of quality? (Mark one)

   - ☐ None
   - ☐ Up to 8 contact hours
   - ☐ More than 8 but less than 40 contact hours
   - ☐ More than 40 contact hours

3. How much time have you spent teaching, training, or briefing others on the subject of quality? (Mark one)

   - ☐ None
   - ☐ Up to 8 contact hours
   - ☐ More than 8 but less than 40 contact hours
   - ☐ More than 40 contact hours

4. Define the term "metric" as it relates to quality.

(continued on next page)

5. List the titles or names of any Metric Development Methods that you are somewhat familiar with.

6. List the titles or names of any Metric Development Methods that you have used to develop metrics.

7. To your knowledge, which of the following have authored books or papers on the subject of quality? (Mark all that apply)

☐ Nancy Brady	☐ H. James Harrington	☐ Genichi Taguchi
☐ Robert C. Camp	☐ Kaoru Ishikawa	☐ Dorsey J. Talley
☐ Philip B. Crosby	☐ J.M. Juran	☐ Arthur R. Tenner
☐ W. Edwards Deming	☐ Dennis Kinlaw	☐ Hans J. Thamhain
☐ Irving J. DeToro	☐ Imai Masaaki	☐ Carl J. Walt
☐ Armand V. Feigenbaum	☐ Tom Peters	☐ Mary Walton
☐ Andrea Gabor	☐ Sonja A. Seefeldt	☐ Lynda Yates

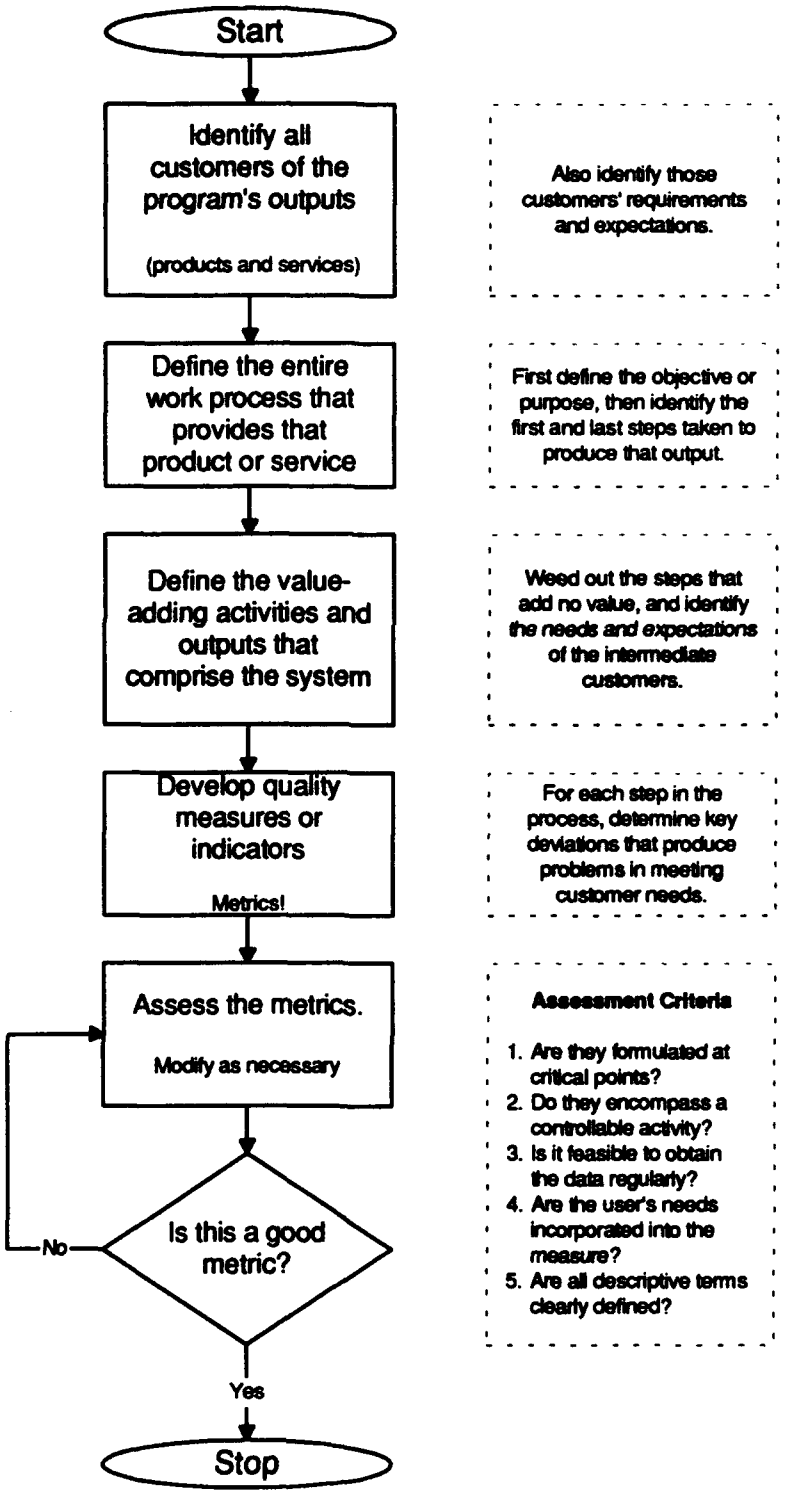
8. Which of the following have authored books or papers on the subject of quality which you have read? (Mark all that apply)

☐ Nancy Brady	☐ H. James Harrington	☐ Genichi Taguchi
☐ Robert C. Camp	☐ Kaoru Ishikawa	☐ Dorsey J. Talley
☐ Philip B. Crosby	☐ J.M. Juran	☐ Arthur R. Tenner
☐ W. Edwards Deming	☐ Dennis Kinlaw	☐ Hans J. Thamhain
☐ Irving J. DeToro	☐ Imai Masaaki	☐ Carl J. Walt
☐ Armand V. Feigenbaum	☐ Tom Peters	☐ Mary Walton
☐ Andrea Gabor	☐ Sonja A. Seefeldt	☐ Lynda Yates

Appendix B. Sample Metric Development Method Questionnaire

OMB Metric Development Method Questionnaire

Please modify the following diagram to ensure it represents the method that you used to develop your metrics.



1. How long has it been since you have used this method to produce a metric?

2. What process, product, or service were you trying to improve?

3. Which steps added value to this method?

4. Which steps did you have to repeat in this method?

5. What information did you require to perform each step?

	Blue Quite a bit	Green Moderate	Yellow Minimal	Red None
6. Of what value was the required information in developing your metric(s)?	☐	☐	☐	☐

7. What was the output product of each step?

8. How did you use the output product of each step?

9. What are the strengths of this method?

10. What are the weaknesses of this method?

11. How would you improve this method?

	Blue Quite a bit	Green Moderately	Yellow Barely	Red Not at all
12. How comfortable were you with using this method?	☐	☐	☐	☐

	Blue Quite a bit	Green Moderately	Yellow Barely	Red Not at all
13. How comfortable are you with the results this method produced?	☐	☐	☐	☐

14. What were the actual metrics produced using this method?

Using the scale provided, please rate the method against each of the following items.

	Blue Quite a bit	Green Moderately	Yellow Barely	Red Not at all
1. How understandable were the steps? (Was the purpose of each step clear?)				
2. How executable were the steps? (Could you actually perform them?)				
3. How efficient was the method? (Did each step add value?)				
4. How logical was the method? (Was the output generated from each step necessary to perform the next step?)				
5. How flexible was the method? (Could the method be logically split into pieces that could be executed over several meetings?)				
6. How sufficient was the method? (Did the method list all the steps necessary to develop a metric?)				
7. How process-oriented was the method? (Did the method force you to focus on a process needing improvement?)				
8. How assessable was the result? (Was a method provided to assess the quality of the metric produced?)				
9. How helpful were the examples provided for each step of the method?				
10. How well did the method force you to focus on the customer of the product or service provided?				

References

Aeronautical Systems Division (ASD). *The ASD Metrics Handbook.* ASD Pamphlet 700-8. Wright-Patterson AFB OH: ASD, 30 April 1992.

Air Force Systems Command (AFSC). *The Metrics Handbook.* Andrews AFB, Washington DC.: HQ AFSC, August 1991.

Advanced Cruise Missile (ACM) System Program Office. *ACM Acquisition Program Tracking System (APTS) Metrics Guide.* Wright-Patterson AFB OH, March 1993.

Basili, Victor R. and H. Dieter Rombach. "Tailoring the Software Process to Project Goals and Environments," *Proceedings of the Ninth International Conference on Software Engineering.* 345-357. New York: IEEE Computer Society Press, 1987.

Buchanan, Lieutenant Stephan. Program Management Project Officer, Aeronautical Systems Center, Aircraft Systems Program Office, Program Management Division (ASC/SDY), Wright-Patterson AFB OH. Personal interview. 27 May 1993.

Caudle, Captain Mark D. *An Analysis of Total Quality Management in Aeronautical Systems Division.* MS Thesis. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991. (AD-A246661).

Camp, Robert C. *Benchmarking - The Search for Industry Best Practices that Lead to Superior Performance.* Milwaukee: ASQC Quality Press, 1989.

Conley, Sonia. Acquisition Management Assistant, Aeronautical Systems Center, Aircraft Systems Program Office, Program Management Division (ASC/SDY), Wright-Patterson AFB OH. Personal interview. 27 May 1993.

Crosby, Philip B. *Quality Is Free.* New York: McGraw-Hill Book Company, 1979.

Deming, W. Edwards. *Out of the Crisis.* Cambridge MA: Massachusetts Institute of Technology Center for Advanced Engineering Study, 1986.

Dierker, Greg. Operations Research Analyst, Aeronautical Systems Center Directorate of Acquisition Logistics (ASC/AL), Wright-Patterson AFB OH. Personal interview. 26 May 1993.

Di Lorenzo, Richard A. Course readings, QMT-084, Quality Management. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, November 1991.

Emory, C. William, and Donald R. Cooper. *Business Research Methods* (Fourth Edition). Boston: Irwin, 1991.

Evans, James R. and William M. Lindsay. *The Management and Control of Quality.* Minneapolis/St. Paul: West Publishing Company, 1993.

Federal Quality Institute (FQI). *Criteria and Scoring Guidelines - The President's Award for Quality and Productivity Improvement.* Washington DC.: Government Printing Office, June 1990.

Galagan, Patricia A. "Training Delivers Results to Federal Express," *Training & Development*, 45: 27-33 (December 1991).

Garofoli, Captain David. Director of Measurement, Aeronautical Systems Center Total Quality Office (ASC/TQ), Wright-Patterson AFB OH. Personal interview. 21 May 1993.

Gibson, James L., John M. Ivancevich, and John H. Donnelly, Jr. *Organizations: Behavior, Structure, Processes* (Seventh Edition). Boston: Irwin, 1991.

Hayes, Captain Robert J., and Captain Lawrence M. Miller. *An Evaluation of Schedule Metrics Used Within Aeronautical Systems Center.* MS Thesis. School of Logistics and Acquisition Management, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1992. (AD-A260113)

Hummel, Brigadier General Dennis K. Vice Commander, Aeronautical Systems Center, Wright-Patterson AFB OH. ASC/CV letter to HQ AFMC/XP. 17 May 1993.

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

Kerr, Steven. "On the Folly of Rewarding A, while Hoping for B," *Academy of Management Journal*, 18: 769-783 (December 1975).

Kinlaw, Dennis C. *Continuous Improvement and Measurement for Total Quality: A Team-Based Approach.* San Diego: Pfeiffer and Co., 1992.

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

Peters, Tom. *Thriving on Chaos: Handbook for a Management Revolution.* New York: Knopf, 1987.

Pytlik, Colonel William F. System Program Director, Aeronautical Systems Center, Advanced Cruise Missile System Program Office (ASC/VC), Wright-Patterson AFB OH. Personal interview. 25 May 1993.

Talley, Dorsey J. *Total Quality Management -- Performance and Cost Measures: The Strategy for Economic Survival.* Milwaukee: ASQC Quality Press, 1991.

Tenner, Arthur R., and Irving J. DeToro. *Total Quality Management: Three Steps to Continuous Improvement.* Reading MA: Addison-Wesley Publishing Company, Inc., 1992.

Thamhain, Hans J. "Benchmarking of Project Management Systems: How to Measure Yourself Against the Best," *Proceedings of the 1991 PMI Seminar/Symposium,* 471-476. Drexel Hill PA: 1991.

Vita

Captain Ken Hamner was born on February 1, 1962 in High Point, North Carolina. He graduated from Ragsdale Senior High School in Jamestown, North Carolina in June 1980 and attended the US Air Force Academy, graduating with a Bachelor of Science degree in Human Factors Engineering in May 1984. Upon graduation, he received a regular commission in the USAF and served his first tour of duty at Eglin AFB, Florida. As an acquisition project officer for the Low Level Laser Guided Bomb Program Office, he managed production deliveries and coordinated new start efforts for the next generation high value target acquisition system. Next, he was reassigned to the Joint Submunition Test and Evaluation Program where he developed the acquisition strategy and coordinated the source selection for the follow-on phase of the program. From there, Captain Hamner entered the Aircraft Maintenance Officer Course at Chanute AFB, IL. Upon graduation in January 1989, he was assigned to the 379th Bombardment Wing at Wurtsmith AFB, MI. First, he served in the Organizational Maintenance Squadron as Assistant Maintenance Supervisor and Tanker Branch OIC. Then he was selected to join the Deputy Commander for Maintenance (DCM) Staff as the Chief of Standardization and Training. In this capacity he led the Wing to its best Maintenance Standardization and Evaluation Inspection performance ever. Finally, he was selected to lead one of the two teams within SAC to test the Maintenance Information Systems Branch concept which established a focal point for small computer use on the DCM Staff. Captain Hamner entered the Graduate School of Logistics and Acquisition Management in May, 1992. He will be assigned to the National Aerospace Plane Joint Program Office upon Graduation.

Permanent Address:
135 Hudson Ave W.
Dayton, OH 45403

Vita

Captain Charles A. La Fleur was born on September 28, 1955 in Albany, New York. He graduated from Ravena-Coeymans-Selkirk Central High School in Ravena, New York, on June 23, 1973. He enlisted in the Air Force on June 26, 1973. Captain La Fleur served as an Electronic Computer and Switching Systems Technician, maintaining the Communications System Segment (CSS) of the NORAD Computer System (427M) at the NORAD Cheyenne Mountain Complex, Colorado. Later, he taught the CSS Computer Maintenance course and was selected for the Airmen Education and Commissioning Program in 1985. He attained the rank of Master Sergeant. In March 1988, he received his Bachelor of Science in Engineering degree from Wright State University, Dayton, Ohio. He attended Officer Training School (OTS) at Lackland AFB, Texas, and was commissioned on September 29, 1988. After OTS, Captain La Fleur was assigned to the Standard Systems Center where he served as an Acquisition Project Officer and a member of the Source Selection Evaluation Team for the Desktop III acquisition. He was reassigned within the Center to a position as a Communications-Computer System Engineering Officer in 1990 where he served as the Test Manager for the Aerospace Safety Automation Program (ASAP) and the Core Automated Maintenance System (CAMS) program. Captain La Fleur entered the Graduate School of Logistics and Acquisition Management in May, 1992. He will be assigned to the Air Force Institute of Technology's Directorate of Communications-Computer Systems upon graduation.

Permanent Address:
5031 Castleview Court
Huber Heights, OH 45424

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
AN EXPLORATORY SURVEY OF METHODS USED TO DEVELOP MEASURES OF PERFORMANCE

5. FUNDING NUMBERS

6. AUTHOR(S)
Kenneth L. Hamner, Captain, USAF
Charles A. La Fleur, Captain USAF

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

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GSM/LAS/93S-6

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Captain Mike Pierce
HQ AFMC/XPD
Wright-Patterson AFB OH 45433

Capt Dave Garofoli
ASC/TQ
Wright-Patterson AFB OH 45433

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)*

Nonmanufacturing organizations are being challenged to provide high-quality products and services to their customers, with an emphasis on continuous process improvement. Measures of performance, referred to as metrics, can be used to foster process improvement. The application of performance measurement to nonmanufacturing processes can be very difficult.

This research explored methods used to develop metrics in nonmanufacturing organizations. Several methods were formally defined in the literature, and the researchers used a two-step screening process to determine the OMB Generic Method was most likely to produce high-quality metrics. The OMB Generic Method was then used to develop metrics. A few other metric development methods were found in use at nonmanufacturing organizations.

The researchers interviewed participants in metric development efforts to determine their satisfaction and to have them identify the strengths and weaknesses of, and recommended improvements to, the metric development methods used. Analysis of participants' responses allowed the researchers to identify the key components of a sound metric development method. Those components were incorporated into a proposed metric development method that was based on the OMB Generic Method, and should be more likely to produce high-quality metrics that will result in continuous process improvement.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Continuous Process Improvement, Measurement, Measures, Metrics, Performance, Quality, Quality Assurance	111
	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/GSM/LAS/93S-6

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