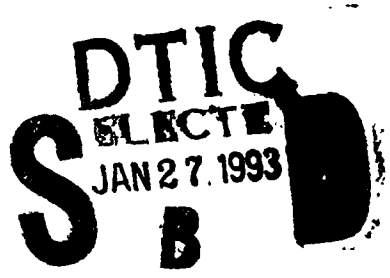
(1)

AFIT/GSM/LSY/92S-12



AN EVALUATION OF SCHEDULE METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

THESIS

Robert J. Hayes, Captain, USAF
and
Lawrence M. Miller, Captain, USAF

AFIT/GSM/LSY/92S-12

Approved for public release; distribution unlimited

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

AFIT/GSM/LSY/92S-12

AN EVALUATION OF SCHEDULE METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

THESIS

Presented to the Faculty of the School of Logistics and Acquisition Management
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Systems Management

Robert J. Hayes, B.S.
Captain, USAF

Lawrence M. Miller, B.S.
Captain, USAF

September 1992

Preface

We decided to focus our thesis work on the concept of metrics - specifically, how metrics are being applied in Air Force acquisition program offices. We narrowed our study to an evaluation of schedule metrics within the Aeronautical Systems Center (ASC), using an expert group to evaluate the extent to which metric-driven behaviors contributed to continuous process improvement. There has been a major push in the Air Force in the last few years to improve quality through metrics. We found that developing effective metrics for acquisition programs is a difficult and complex task. Our main hope for this project was that it could provide useful information to other acquisition program offices challenged to develop their own internal metrics.

We've had a great deal of help during the course of our research. We'd like to thank Capt Ken Moen and Ms. Janet Peasant at Armstrong Laboratory for their help and for allowing us to use the human resource laboratory. We'd also like to thank Capt Randy Kosinski at HQ ASC\CCX for his help in obtaining the metrics. Our advisors, Majors Kevin Grant and Wendell Simpson, provided expert guidance and demonstrated remarkable patience over our long, frustrating voyage. A final thanks goes to our wives, Lisa and Margie, and our children without whose patience, understanding, and support this would not have been possible.

Table of Contents

	Page
Preface	ii
List of Figures	v
List of Tables	x
Abstract	xi
I. Introduction	1-1
Purpose and Overview	1-3
Problem Statement	1-5
Research Questions	1-5
Scope	1-6
Summary	1-7
II. Literature Review	2-1
Chapter Overview	2-1
Importance of Metrics	2-1
Attributes of a Good Metric	2-9
Meaningfulness	2-9
Linked to Goals and Objectives	2-10
Simple, Understandable, Logical, and Repeatable	2-11
Shows a Trend	2-12
Unambiguously Defined	2-12
Economical to Collect	2-12
Timely	2-13
Drives the Appropriate Behavior	2-13
Summary	2-14
III. Methodology	3-1
Metrics Collection	3-1
Mail Survey	3-2
Metric Presentation	3-3
Metric Selection	3-4
Metric Taxonomy	3-5
Metric Selection Criteria	3-6
Selected Metrics	3-7
Metric Evaluation	3-10
Evaluation Criteria	3-10
Evaluation Method	3-10
Group Evaluation Process Development	3-14
Group Evaluation Process Execution	3-16
Results of the Evaluation	3-23
Summary	3-28

Page

IV. Metrics Evaluation and Analysis 4-1

Metric 1 4-1
Rating Summary for Metric 1 4-3
Evaluation Group Comments 4-3
Analysis 4-4
Metric 2 4-7
Rating Summary for Metric 2 4-8
Evaluation Group Comments 4-8
Analysis 4-10
Metric 3 4-11
Rating Summary for Metric 3 4-12
Evaluation Group Comments 4-12
Analysis 4-14
Metric 4 4-15
Rating Summary for Metric 4 4-16
Evaluation Group Comments 4-17
Analysis 4-18
Metric 5 4-20
Rating Summary for Metric 5 4-22
Evaluation Group Comments 4-22
Analysis 4-23
Metric 6 4-25
Rating Summary for Metric 6 4-26
Evaluation Group Comments 4-27
Analysis 4-28
Metric 7 4-30
Rating Summary for Metric 7 4-31
Evaluation Group Comments 4-31
Analysis 4-31

V. Conclusions and Recommendations 5-1

Conclusions 5-1
Recommendations for Further Research 5-6
Summary 5-9

Appendix A. Individual Behavior Graphs A-1

Appendix B. Evaluation Session Help Aids B-1

Bibliography Bib.1

Vita Vita-1

List of Figures

Figure		Page
3-1.	Line Graph Representation	3-4
3-2.	Bar Graph Representation	3-5
3-3.	Area Graph Representation	3-6
3-4.	Topic Commenter Screen	3-18
3-5.	Alternative Evaluator Screen	3-20
3-6.	Cloud Chart for Single Behavior	3-22
3-7.	Quadrant One Behaviors	3-24
3-8.	Quadrant Two Behaviors	3-25
3-9.	Quadrant Three Behaviors	3-26
3-10.	Quadrant Four Behaviors	3-27
4-1.	Results of Metric 1	4-3
4-2.	Results of Metric 2	4-9
4-3.	Results of Metric 3	4-13
4-4.	Results of Metric 4	4-17
4-5.	Results of Metric 5	4-22
4-6.	Results of Metric 6	4-27
4-7.	Results of Metric 7	4-32
A-1.	Metric One, Behavior One	A-1
A-2.	Metric One, Behavior Two	A-2
A-3.	Metric One, Behavior Three	A-3
A-4.	Metric One, Behavior Four	A-4
A-5.	Metric One, Behavior Five	A-5
A-6.	Metric One, Behavior Six	A-6
A-7.	Metric One, Behavior Seven	A-7

Figure Page

A-8. Metric One, Behavior Eight A-8

A-9. Metric One, Behavior Nine A-9

A-10. Metric One, Behavior Ten A-10

A-11. Metric One, Behavior Eleven A-11

A-12. Metric One, Behavior Twelve A-12

A-13. Metric One, Behavior Thirteen A-13

A-14. Metric One, Behavior Fourteen A-14

A-15. Metric One, Behavior Fifteen A-15

A-16. Metric Two, Behavior One A-16

A-17. Metric Two, Behavior Two A-17

A-18. Metric Two, Behavior Three A-18

A-19. Metric Two, Behavior Four A-19

A-20. Metric Two, Behavior Five A-20

A-21. Metric Two, Behavior Six A-21

A-22. Metric Two, Behavior Seven A-22

A-23. Metric Two, Behavior Eight A-23

A-24. Metric Two, Behavior Nine A-24

A-25. Metric Three, Behavior One A-25

A-26. Metric Three, Behavior Two A-26

A-27. Metric Three, Behavior Three A-27

A-28. Metric Three, Behavior Four A-28

A-29. Metric Three, Behavior Five A-29

A-30. Metric Three, Behavior Six A-30

A-31. Metric Three, Behavior Seven A-31

A-32. Metric Three, Behavior Eight A-32

Figure Page

A-33. Metric Three, Behavior Nine A-33

A-34. Metric Three, Behavior Ten A-34

A-35. Metric Three, Behavior Eleven A-35

A-36. Metric Three, Behavior Twelve A-36

A-37. Metric Three, Behavior Thirteen A-37

A-38. Metric Three, Behavior Fourteen A-38

A-39. Metric Three, Behavior Fifteen A-39

A-40. Metric Four, Behavior One A-40

A-41. Metric Four, Behavior Two A-41

A-42. Metric Four, Behavior Three A-42

A-43. Metric Four, Behavior Four A-43

A-44. Metric Four, Behavior Five A-44

A-45. Metric Four, Behavior Six A-45

A-46. Metric Four, Behavior Seven A-46

A-47. Metric Four, Behavior Eight A-47

A-48. Metric Four, Behavior Nine A-48

A-49. Metric Four, Behavior Ten A-49

A-50. Metric Four, Behavior Eleven A-50

A-51. Metric Four, Behavior Twelve A-51

A-52. Metric Four, Behavior Thirteen A-52

A-53. Metric Four, Behavior Fourteen A-53

A-54. Metric Four, Behavior Fifteen A-54

A-55. Metric Four, Behavior Sixteen A-55

A-56. Metric Five, Behavior One A-56

A-57. Metric Five, Behavior Two A-57

A-58. Metric Five, Behavior Three A-58

Figure Page

A-59. Metric Five, Behavior Four A-59

A-60. Metric Five, Behavior Five A-60

A-61. Metric Five, Behavior Six A-61

A-62. Metric Five, Behavior Seven A-62

A-63. Metric Five, Behavior Eight A-63

A-64. Metric Five, Behavior Nine A-64

A-65. Metric Five, Behavior Ten A-65

A-66. Metric Five, Behavior Eleven A-66

A-67. Metric Five, Behavior Twelve A-67

A-68. Metric Five, Behavior Thirteen A-68

A-69. Metric Five, Behavior Fourteen A-69

A-70. Metric Five, Behavior Fifteen A-70

A-71. Metric Five, Behavior Sixteen A-71

A-72. Metric Five, Behavior Seventeen A-72

A-73. Metric Five, Behavior Eighteen A-73

A-74. Metric Five, Behavior Nineteen A-74

A-75. Metric Five, Behavior Twenty A-75

A-76. Metric Five, Behavior Twenty-one A-76

A-77. Metric Six, Behavior One A-77

A-78. Metric Six, Behavior Two A-78

A-79. Metric Six, Behavior Three A-79

A-80. Metric Six, Behavior Four A-80

A-81. Metric Six, Behavior Five A-81

A-82. Metric Six, Behavior Six A-82

A-83. Metric Six, Behavior Seven A-83

A-84. Metric Six, Behavior Eight A-84

Figure Page

A-85. Metric Six, Behavior Nine A-85

A-86. Metric Six, Behavior Ten A-86

A-87. Metric Six, Behavior Eleven A-87

A-88. Metric Six, Behavior Twelve A-88

A-89. Metric Six, Behavior Thirteen A-89

A-90. Metric Six, Behavior Fourteen A-90

A-91. Metric Seven, Behavior One A-91

A-92. Metric Seven, Behavior Two A-92

A-93. Metric Seven, Behavior Three A-93

A-94. Metric Seven, Behavior Four A-94

A-95. Metric Seven, Behavior Five A-95

A-96. Metric Seven, Behavior Six A-96

A-97. Metric Seven, Behavior Seven A-97

A-98. Metric Seven, Behavior Eight A-98

A-99. Metric Seven, Behavior Nine A-99

A-100. Metric Seven, Behavior Ten A-100

A-101. Metric Seven, Behavior Eleven A-101

A-102. Metric Seven, Behavior Twelve A-102

A-103. Metric Seven, Behavior Thirteen A-103

List of Tables

Table		Page
3-1.	Research Methodology Summary	3-28
4-1.	List of metrics	4-1
4-2.	Likely Behaviors Resulting From Metric 1 - Number of Product Delivered vs. Number of Product Required	4-2
4-3.	Likely Behaviors Resulting From Metric 2 - Average Number of Days to Respond to Service Report	4-7
4-4.	Likely Behaviors Resulting From Metric 3 - Number of UCAs Definitized Within 180 Days vs. Number Definitized After 180 Days	4-11
4-5.	Likely Behaviors Resulting From Metric 4 - Number of Project Baseline Breaches	4-15
4-6.	Likely Behaviors Resulting From Metric 5 - Average Variance in Days Between Scheduled Release of RFP and Actual Release	4-20
4-7.	Likely Behaviors Resulting From Metric 6 - Number of Contractor Data Submittals Due vs. Received	4-25
4-8.	Likely Behaviors Resulting From Metric 7 - Number of GFE Deliveries Provided vs. Number Required	4-30

Abstract

This study focused on a selected group of schedule metrics in use at Aeronautical Systems Center (ASC) acquisition program offices. Over 300 metrics were collected from the System Program Offices (SPOs). The metrics data was sorted into the categories of cost, schedule, and performance. In order to narrow the scope of the project, the team focussed on schedule metrics. Seven of the most common schedule metrics were selected for evaluation by a group of five experienced acquisition professionals. Using a Group Support System at Armstrong Laboratory, Wright-Patterson AFB, the group was asked to determine the likely behaviors driven by the metrics being addressed by the study. Next, they were asked to rate the metrics according to how well the metric-driven behaviors contributed to continuous improvement. Results of the study showed a wide spread of behaviors, both positive and negative, that would likely be driven by the metrics. Most of the metrics rated were found to need improvement in terms of influencing behaviors that would lead to continuous process improvement. However, many good metric-driven behaviors were identified which could prove helpful to program offices undergoing the challenge of developing their own internal metrics.

AN EVALUATION OF SCHEDULE METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

I. Introduction

Due to the perception that the United States is losing its position as the world's dominant economic power, managerial experts have been increasingly focusing their attention on the areas of competitiveness and quality. Statistics only confirm this perception.

Of the world's 35 largest banks, only one is from the United States. Among the world's ten largest corporations, only three are from this country. The United States has fallen from world leadership in per capita income and individual productivity (34:2). In addition to the examples mentioned, there are numerous others that illustrate the United States' slide in competitiveness and quality in the world market.

The Department of Defense (DOD) has a special interest in the erosion of quality and productivity in defense related businesses. The United States depends on technological superiority to offset the numerical advantage of its potential adversaries. In addition, declining budgets in the post-Soviet Union era demand that

productivity and quality be emphasized more than ever to get the most out of scarce defense dollars (10:xi). These concerns have led to a major effort to improve productivity and quality throughout industry and government within the United States. One of the primary methods being pursued to achieve these goals is the concept of Total Quality Management (TQM). TQM, as defined by the Office of the Secretary of Defense, is:

> both a philosophy and a set of guiding principles that represent the foundation of a continuously improving organization. TQM is the application of quantitative methods and human resources to improve the material and services supplied to an organization, all the processes within an organization, and the degree to which the needs of the customer are met, now and in the future. (12:1)

More succinctly, "TQM is a participative management style which focuses on satisfying customer expectations by continually improving the way business is conducted" (34:viii).

According to the office of the Deputy Undersecretary of Defense for TQM, there are four pillars of TQM:

1. Organizations must be customer driven and customer responsive.
2. The customer defines quality.
3. Process improvement never ends.
4. People who run the trains know the tracks best (13:14).

Other lists of tenets of TQM exist as well. However, all the lists can essentially be synthesized into four fundamental components:

1. Leadership
2. Complete customer focus
3. Continuous process improvement
4. Empowerment of employees (15:34).

Purpose and Overview

While it may be necessary to implement all four of the fundamental components of TQM to realize success, this study will focus mainly on the third guideline -- continuous process improvement. Specifically, the central objective is to examine the concept of measurement. Measurement is one of the most important ways that continuous process improvement can be achieved.

> Processes operated without measurement are processes about which very little is known. Conversely, if inputs and outputs can be measured and expressed in numbers, then something is known about the process, and control is possible. (29:208)

Stated another way, what can be controlled has the possibility of being improved. Improvement isn't possible unless the status of the process is known. Knowledge of the process is impossible without some system of measurement.

Measurement alone is not sufficient for process improvement. The measurements must be meaningful. As the

Metrics Handbook prepared by Air Force Systems Command (AFSC)[1] states:

> Metrics are nothing more than meaningful measures. For a measure to be meaningful, however, it must present data that allow us to take action. It must be customer oriented and support the meeting of our organizational goals and objectives. Metrics foster process understanding and motivate action to continually improve the way we do business. This is distinguished from measurement, in that, measurement does not necessarily result in process improvement. Good metrics always will. (4:1-1)

It appears evident that the concept of metrics is critical to the process of continuous improvement.

One United States Air Force (USAF) organization that is making an active commitment toward implementing TQM and the resultant concept of metrics is the Aeronautical Systems Center (ASC). ASC, a constituent product center of Air Force Materiel Command (AFMC), manages the development and acquisition of aeronautical systems and related equipment (30:16). ASC has responded to the AFSC Commander's guidance on metrics:

> Measurement is a fundamental part of good management. Metrics are invaluable from both a program management and a process improvement perspective. As such, we need metrics to be an integral part of daily operations throughout Air Force Systems Command. Measuring processes provides the basis for appropriate management action(s) to identify opportunities for constructive changes and continuous improvement. Metrics allow us to baseline where we are, identify the impediments to the process, and track

---

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

> the impact of management actions on processes and other process changes. (2:2.4)

This response has come in the form of the Acquisition Program Tracking System (APTS). "APTS was designed to automate the collection, output, and analysis of ASC metrics and consequently track progress toward meeting ASC objectives" (1:21). This information is primarily used by the ASC Commander, System Program Office (SPO) directors, and senior level management (1:1). SPOs are also encouraged to develop and track their own metrics as well. These metrics can be aggregated into APTS for use as product center metrics or they can be used to improve processes within the SPO itself (1:19). Regardless of the level of reporting, the purpose of ASC metrics is to drive appropriate behaviors which will result in continuous improvement (1:21).

Problem Statement

ASC has devoted considerable effort to developing and implementing APTS. However, most of the SPOs have developed their internal metrics on their own, with much less guidance or expertise (6,23). Because these internal metrics have the potential to result in substantial improvement to SPO processes, this study will focus on an evaluation of the quality of these metrics. The objective of this study is to identify the behaviors which will likely result from the use

of common schedule metrics and to evaluate the extent to which these behaviors will lead to continuous improvement.

Research Questions

The following investigative questions were used to accomplish the objective of the research:

1. What are the criteria of a good metric?
2. What are the internal schedule metrics currently in use within ASC SPOs?
3. What are the possible behaviors that are driven by these schedule metrics?
4. Do the behaviors driven by the schedule metrics lead to continuous improvement?

Scope

This study concentrated on those internal metrics used within the SPOs of ASC at Wright-Patterson AFB, Ohio. Although the schedule metrics in question may feed into APTS, no APTS metrics were evaluated.

Even though the study was limited to ASC, the results may be applicable to the other acquisition centers in AFMC. The activities at Electronic Systems Center (ESC), Space and Missile Center (SMC), and Human Systems Center (HSC) all focus on the acquisition of weapon systems. In addition, the personnel at each center have received the same basic training and utilize the same processes to develop and acquire new systems.

The number of metrics evaluated is a final limitation of the study. Because of the large number and variety of metrics currently in use, only a subset was chosen for evaluation. The metrics were segregated into major categories of cost, schedule, and performance. Upon review, the subset of schedule metrics proved to be easiest to categorize and present in a generic format and were found to include examples applicable to most of the SPOs involved. The generalizations made concerning these metrics may not apply to other metrics in use within ASC, especially performance metrics, because of the SPO-unique nature of this category.

Summary

This chapter provided a brief introduction to one of the most critical aspects of successful implementation of TQM, measurement and metrics. It also provided a brief overview of metrics use within ASC. In addition, it discussed the research problem and investigative questions which form the basis of the thesis. The chapter concluded with a discussion concerning the scope and limitations of the research.

Chapter two provides an in-depth review of the concept of measurement and metrics. Chapter three focuses on the methodology used to execute the research. The fourth chapter presents the findings and results of the research.

Finally, chapter five provides the researchers' conclusions and recommendations for further research in the area of metrics.

II. Literature Review

The following passage provides an important insight on the subject of measurement:

> You cannot manage what you cannot measure. You cannot measure what you cannot operationally define. You cannot operationally define what you do not understand...You will not succeed if you do not manage. (33:74)

Chapter Overview

Building on the introduction to the concept of measurement and metrics in the preceding chapter, this section will provide a summary of the current literature in this area. Specifically, it will focus on why measurement and metrics are important in any system emphasizing continuous improvement. In addition, the chapter will address some attributes of an effective metric gleaned from the literature review.

Importance of Metrics

Dr. W. Edwards Deming, one of the most renowned experts in the Total Quality Management (TQM) arena, has developed a list of steps (points) necessary to improve quality and productivity. These steps are referred to as Deming's Fourteen points. Point number five specifically addresses continuous improvement: "Improve constantly and forever the system of production and service, to improve quality and productivity, and thus constantly decrease costs" (11:23). Deming places great emphasis on the use of statistical

control charts in order to achieve stability. Use of these control charts can help the manager determine whether variation in a process is attributable to a specific worker, machine, or local condition. Deming refers to this as "special cause."

Special causes are not inherent to the process. Conversely, common causes are. Deming defines common causes of variation as those that are inherent to the process itself. Common causes can only be dealt with by management. His inherent assumption is that workers have no control over the process being used (11:310-314). This delineation is important if continuous improvement is the goal. "Eliminating special causes of variation does nothing to improve a process; it merely puts the process back where it should have been anyway" (7:3-44). Once special causes have been eliminated, the only way to continually improve is to eliminate common causes of variation. Stable control charts indicate that any variation left is probably due to a common cause. "Improvement of the process can be pushed effectively, once statistical control is achieved and maintained" (11:321). Conversely, action taken to change a stable process due to a defect (i.e special cause of variation), will only destabilize a previously stable process (11:20). Therefore, it is imperative to know whether the process is stable or not.

Improvement of the process then, entails some type of adjustment to the system and an observation of the results of the change to the system. This is where the concept of measurement becomes so important.

"Measurement is feedback. It provides information needed to take corrective action if required (9:1)." In order to make any change, a manager must know what the process is doing and how it is being altered. Without any meaningful measure, it would be impossible to make effective adjustments. If there doesn't appear to be a starting point, how can an effective adjustment be made?

> If you don't measure, how will you know if you are:
> 1. Getting the job done within specifications?
> 2. Meeting your long range needs?
> 3. Improving fast enough?
>
> You will know only if you have properly designed and executed measurement and evaluation systems. (33:74)

In his book Quality is Free, Philip Crosby addresses the requirement for measurement: "Quality is free, but no one is ever going to know it if there isn't some sort of agreed-on system of measurement" (8:121).

Similar to Deming, Crosby lists fourteen steps which need to be followed to improve quality. Step three deals with quality measurement. The purpose of this step is "to provide a display of current and potential nonconformance problems in a manner that permits objective evaluation and corrective action" (8:199). Crosby defines quality as

conformance to requirements. When defined in this manner, quality becomes measurable (8:17). Measurement, again, appears to play a key role in improving quality. Before any improvement can take place, the current status of quality must be determined. Measurements need to be developed then used to identify areas for improvement (7:3-58).

Joseph M. Juran, another expert in the area of quality, has developed what is called the Juran Trilogy. Juran views the trilogy as a blueprint for quality management. The components of this plan include: quality planning, quality control, and quality improvement. Measurement is heavily stressed in the last two components. Quality control consists of evaluating actual performance, comparing actual performance to previously established goals, and acting on the differences. Quality improvement is achieved by establishing goals, and reviewing progress toward meeting those goals.

One stumbling block to improving quality involves the measurement of parameters of quality improvement. Juran's contention is that new methods of measurement may need to be developed. He calls these "new metrics" (21:69). According to Juran, good quality is not attainable without precise communication among all involved. This precision can best be obtained by stating things numerically. This requires a system of measurement (21:153).

Many government agencies, on the local, state, and federal levels, have undertaken some form of TQM (28:198).

TQM efforts have become especially pervasive throughout the federal sector. The Office of Management and Budget (OMB) spearheads this effort. OMB lists ten elements as keys to an effective implementation of a TQM system in the federal sector. Three of the elements address measurement to some extent (28:198).

The first element focuses on the identification of quality goals and the development of an annual improvement plan. This is consistent with Juran and Crosby's proposals; however, Deming would disagree. According to Deming, the quality attainable is set by the system itself. Goals are not necessary (11:76).

The second element focuses on developing these productivity and quality goals with a special consideration toward the user (28:198). In other words, the goals should be meaningful to those that will use them.

The third and final element addresses the results of the measurement. The results of the measurements need to be compared to the improvement plan in order to get any value out of the process. In this way, the results can be used to hold managers and employees accountable for their actions and decisions (28:198). This is a point with which all three of the above mentioned would disagree. Deming's eighth point deals with eliminating fear. Fearful employees who are afraid to speak up, ask questions, or innovate do not perform at their peak capability (11:59). Management should first look to the process involved to determine if it

is in control. If it is, then special causes can usually be found to be the source of problems. Only then should individuals being singled out. Even then, the intent should not be to punish but to improve (11:62).

Similarly, Crosby feels that management is the source of most quality problems because management has created the system which the workers must use. Holding people responsible for problems they have no control over will lead to frustration. Employees must understand that measurement will be used to identify areas for improvement, not to place blame. The key is that management must accept responsibility for defective processes and encourage workers to participate in process improvement (8:16,112-119).

Juran feels that, although goals are important to attaining quality improvement, it is management that is responsible for making changes when goals are not achieved. Specifically, management must "establish the infrastructure needed to secure annual quality improvement" (21:21). The primary source of improvement is management. Holding workers accountable for the responsibilities of management will not improve quality.

In a study of public sector productivity, Peter Drucker identified several barriers to improvement, two of which speak directly to measurement deficiencies: a lack of clear performance targets and the lack of evaluation (14:103-106). In other words, when there is no baseline

from which to measure performance and also a lack of a measurement mechanism, problems can arise (28:198).

One of the results of the emphasis on continuous improvement was the establishment of The President's Quality Improvement Prototype (QIP) Award. This award recognizes federal organizations that have improved the efficiency, quality, and timeliness of their products and services. One of the scoring categories specifically addresses measurement and analysis. The emphasis is toward collecting data that will help the organization improve its services or products (17:21).

The Malcolm Baldridge National Quality Award is another award that emphasizes measurement as a tool in achieving continuous improvement. Since Congress created the award in 1987 to recognize companies producing world-class goods and services, it has become the standard of excellence for American business (26:101). The Baldridge Award criteria consist of eight essential factors by which judges evaluate candidate companies. Two of these criteria relate directly to measurement: "a plan to keep improving all operations continuously and a system for measuring these improvements accurately" (26:108). Points are awarded in six categories carrying specific point totals. The second-highest ranking category, Quality Results, looks specifically at quality levels and improvements based on measures. The measures are driven by the emphasis on customer satisfaction -- the top-ranked category (31:39).

A more succinct description of TQM and the role of measurement in continuous improvement is provided by the Training Systems Program office, a subordinate unit of ASC located at Wright-Patterson AFB, Ohio. Measurement is considered to be one of the seven quality fundamentals within the TQM effort in this program (25:39). The benefit of measuring has been found to be substantial within this program:

> [Measurement] enables us to focus on specific problem areas like office communications that affect our internal customers; quantifies an intangible; gets you closer to the unbiased organization pulse; sends the right message - we care about what people think. (25:44)

ASC has recognized the importance of metrics and the concept of measurement:

> Good metrics facilitate improvement. Metrics help us understand our many processes (and their capabilities) so we can continually improve them. They can identify processes, or portions of processes, which are "broken" and provide insight to the reasons for problematic processes. Metrics will help identify negative and positive trends...Armed with the knowledge of what causes problems, we can devise a solution and make intelligent decisions to change our processes and share successes across the organization. (1:7)

The above examples clearly indicate the importance placed on effective measurements, or metrics, in any system emphasizing continuous improvement such as TQM. Now that the importance of metrics has been established, the focus will shift to the attributes of a good metric.

Attributes of a Good Metric

AFSC has developed a list of attributes of a good metric to use as a framework for developing good metrics. Specifically, attributes of a good metric include the following.

1. It is acceptable as meaningful to the customer.

2. It tells how well organizational goals and objectives are being met through processes and tasks.

3. It is simple, understandable, logical, and repeatable.

4. It shows a trend.

5. It is unambiguously defined.

6. Its data is economical to collect.

7. It is timely.

8. It drives the appropriate action. (4:2-1)

A discussion of each criterion follows.

Meaningfulness. Meaningful to the customer relates to the heart of the TQM approach to continuous improvement. Every organization, no matter how small, has customers who are its reason for existence. These customers, whether internal or external to the organization, must be considered when metrics are being developed. Unless the metrics relate to improvements from the customer's viewpoint there is no sense in using that measure. Meaningfulness also applies to those who must collect the data to build the metric as well as those who must use the information for evaluation purposes. In addition, those workers being measured by the

metric must understand how they are being measured and why. Metrics will be far more effective if they are accepted by the people involved in the process being measured (29:210). This implies a team focus to the development of the metrics as well as their use (33:79). Additionally, metrics should not be applied to processes outside the responsibility of the team. "You can only change something if it is within your authority" (9:2). Finally, meaningfulness to all involved is possible only if everyone is kept informed. Continuous effort must be made to ensure that everyone knows what is being measured, why it is being measured, and what actions are being taken as a result (9:2). Crosby nicely summarizes the importance of meaningfulness: "Quality measurement is only effective when it is done in a manner that produces information people can understand and use" (8:199).

Linked to Goals and Objectives. Metrics should have broad applicability to the goals and objectives that they support. They should help provide answers to such questions as: Is quality improving? Are we competitive? How can we improve? (22:76) A metric linked to goals and objectives is one that is most helpful to the recipient in answering these types of questions. Conversely, "if you are trying to accomplish A but you are measuring B, A will not happen" (9:2). In other words, to achieve a goal, a metric must provide information that can readily be used to move toward that goal. Finally, metrics should be used as means of

improvement, not as a means of determining who is at fault. A sure way to get unreliable or even falsified information is to use it to punish people (9:2).

Simple, Understandable, Logical, and Repeatable. Simplicity in this case has two parts. First, metrics should not be complex. In other words, they should be understandable to all those involved in the measurement process (8:199). Secondly, metrics can be rather simple to use if a ready-made system for collection already exists. This is usually the case on a technological level (i.e. machine performance). However, in a managerial atmosphere, there may not be a method of measurement in existence. If this is the case, it may be necessary to create one. Juran believes that creating a system of measurement on the managerial level should not be a problem, although he cautions against vagueness and unnecessary complexity to ensure understandability. A lack of understanding can lead to suspicions on the part of those who may not understand the purpose clearly (22:78). Crosby, on the other hand, believes developing measurements is often difficult, particularly in service-type industries. He suggests that the entire team be involved in development of metrics to ensure everyone understands (8:201). Closely related to understanding is the criterion of being logical. A metric must make sense to those using it. "People are much more likely to accept them if they understand why they are important" (9:5). A common problem in this area occurs when

excessive metrics are being used. It is easy to measure just for the sake of measuring. The emphasis should be on the quality of the information being collected, not the quantity. "Don't try to measure everything. A few good measures is the best policy" (9:2).

Shows a Trend. A one-time measure cannot provide any indication whether or not improvement is taking place (4:2-2). A metric should be repeatable over time in order to provide trend information. Management can then use trend results to determine if improvement is taking place. Using Deming's point about improving constantly and forever, it follows that metrics must provide trend information (11:49-52).

Unambiguously Defined. This criterion is closely related to understandability. Understandability relates to why a metric is being used, while an unambiguous definition describes what is being measured. A metric should be susceptible to uniform interpretation by those involved with its use. "Identical numbers can nevertheless result in widely different interpretations. What is critical is whether the units of measure have been defined with adequate precision" (22:77). At the managerial level, many measures may involve words that lack standardized meanings. It is imperative that all involved know what is being measured and how the process occurs.

Economical to Collect. This criterion asks the question: Is the cost of performing the measurement more

than the value received from obtaining the metric information? If the answer is yes, then the metric should probably not be used. However, it may be that the metric itself provides useful information but the precision of measurement being used is too stringent. Relaxing the precision of the measurement may result in the metric becoming economical to use (22:78).

Timely. The following clearly states the need for timely information:

> Measure as close to the activity as is reasonable. Measurements made close to the activity are the easiest to react to in a timely and effective way. Measurements made too high in an organization are difficult to react to because of traceability and timeliness. (9:2)

Late information is useless. Worse yet, it's a waste of time, money, and effort.

Drives the Appropriate Behavior. The Metrics Handbook considers this the most important criterion of a good metric (4:2-1). Essentially, this criterion focuses on whether or not the metric will lead to actions which will result in improvement (29:209). The metric should relate directly to performance and should provide a means of evaluating and motivating improvement (33:79). One important point to note is that metrics should be used in concert with one another as opposed to singularly. The action that is driven by a single metric may lead to improvement in one area while some other part of the process or organization may suffer. Additionally, the metric should be used for improvement

only. As previously stated, using metrics to determine who is at fault, with the intention of punishing the offending party, will destroy any possible benefit of the metric (9:2).

Measurement, by itself may not be sufficient to insure that actions are taken toward improving continuously. If the measures do not provide useful information concerning the process in question, then the measurement effort has been wasted. A measurement can be valid from a procedural viewpoint, but useless from an improvement standpoint (2:2.1).

> For a measure to be meaningful it must present data that allows us to take appropriate actions. It should be customer oriented and should foster process understanding, thereby motivating action to continually improve the way we do business. (2:2.1)

Summary

The concept of measurement, or metrics, is pervasive throughout any system emphasizing continuous improvement. Without measurement, improvement is wishful thinking. The national push for improved productivity and quality through programs such as TQM has provided the impetus for the design of many measurement systems. Criteria vary from program to program but have one fundamental purpose: to drive behavior that will lead to continuous improvement.

III. METHODOLOGY

In order to adequately address the research questions posed in chapter one, the research team members developed a five step methodology. First, the team conducted a literature review to explore the importance of measurement to management systems emphasizing continuous improvement and to determine the attributes of an effective metric. Second, the research team identified and selected the Aeronautical Systems Center (ASC) System Program Offices (SPOs) that were considered to be the "front runners" in the internal SPO metric development process. Third, the SPOs identified in step two were asked to provide their internal metrics data. Fourth, a metrics evaluation group was identified. And finally, the identified group evaluated the metrics through a Group Support System (GSS) provided by Armstrong Laboratory (AFMC/AL/HRG).

The objective of the research is to determine whether the selected metrics drive behaviors that result in continuous improvement. The methodology used to answer each research question will now be addressed.

Metrics Collection

ASC is the acknowledged AFMC leader in the implementation of Total Quality (3:iii). Therefore, a logical starting point in the collection of candidate

metrics was ASC. Realizing that an evaluation of all ASC metrics would be an impossible undertaking due to the sheer number involved, the focus was narrowed to include only those metrics internal to the System Program Offices (SPOs). The SPO is where all acquisition functions come together to "define, acquire, and deliver a system or combination of systems" (3:i). Focusing on the SPOs allowed an evaluation of factors that directly influence the quality of an end product.

Mail Survey. The research team used the mail to collect the metrics. One of the main advantages of collecting data by mail is the fact that respondents can take more time to collect facts. This can increase the quality of the data (16:333). Based on recommendations from the office of primary responsibility for the development of APTS (ASC/CYN) and the office tasked with implementing the use of APTS (ASC/CCX), five SPOs were identified as being "out in front" in the metrics development process. These offices included ASC/RW (Reconnaissance), YT (Simulators), SD (Aircraft), VF (F-15), and VL (Low Altitude Navigation and Targeting Infrared for Night) (LANTIRN)) (6,23). With the assistance of ASC/CCX, ASC/CC signed a letter directing these SPOs to provide their internal metrics data to the research team. As a result, over 300 metrics were collected. As previously stated, these metrics were not APTS metrics but those internal to each SPO. APTS metrics have received extensive management attention, both from a

development and analysis perspective. In contrast, SPO-internal metrics have received far less attention. For this reason, the SPO-internal metrics were chosen for analysis (23). In addition, these metrics are the tools that are actually used by the SPOs for internal improvement.

Metric Presentation. Review of the metrics received in response to the mail request revealed the SPOs were using a variety of metric presentation schemes. Evaluating the metrics in the manner in which they were received could have biased the research. The graphical representations may have induced the evaluators to focus on the appearance of the metric rather than the substance and intent of the metric. For this reason, the most common schedule metrics were described "textually" based on the specific performance characteristic being measured. An example of this would be "Units delivered versus units required on a monthly basis." Figure 3-1 is a line graph representation showing units required versus delivered. It is relatively simple to determine how many deliveries were late. Figure 3-2 is a bar graph which shows percentages delinquent and on-time. Instead of absolute numbers it provides information on the relative success of meeting delivery schedules. Figure 3-3 is similar to Figure 3-1 in that it shows the number of late deliveries but uses an area graph as opposed to a line graph. This provides information on the cumulative numbers of late deliveries. Regardless of the presentation method,

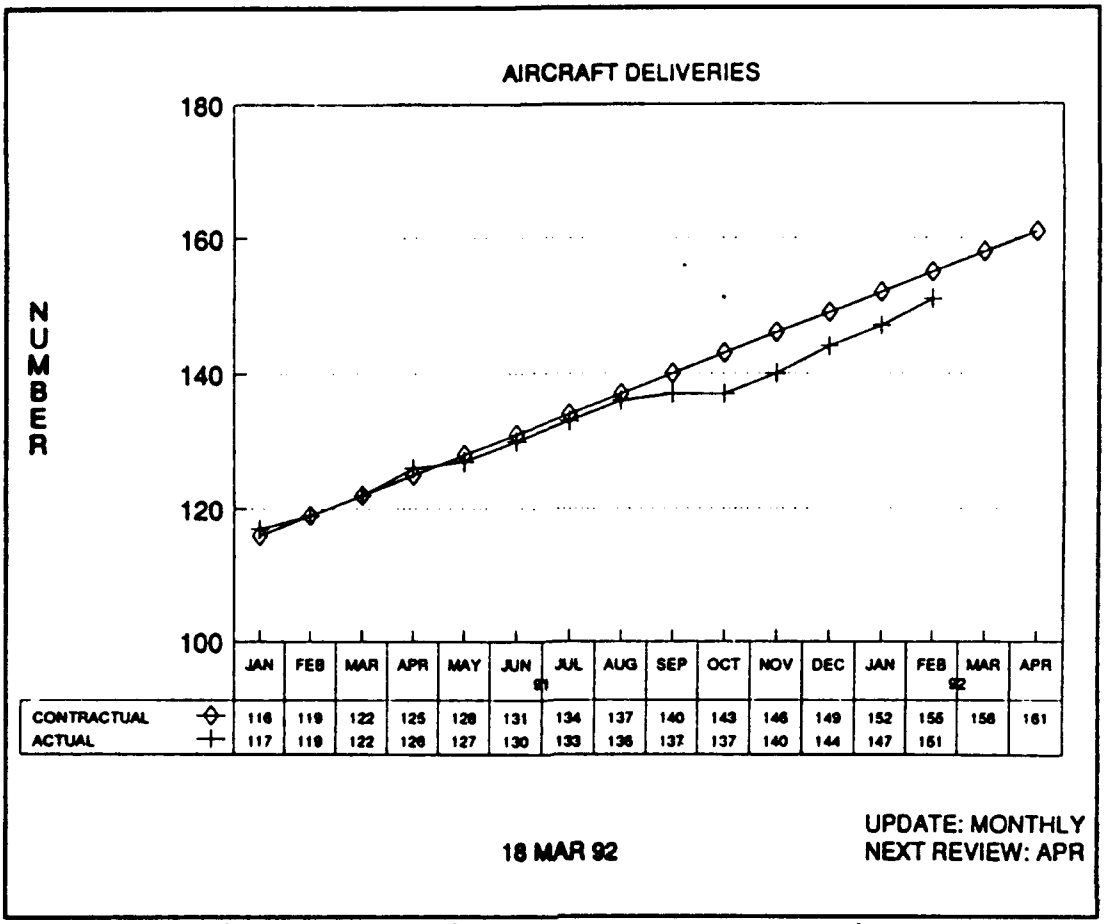


Figure 3-1. Line Graph Representation

each presentation is trying to convey the same information: a comparison of product required to be delivered to product actually delivered. Because individuals have different preferences for methods of presentation, this factor was removed from the evaluation process to ensure a focus on the metric, not the presentation style.

Metric Selection

Upon receiving the metrics from the SPOs, a two-step process was employed to select metrics for evaluation. The

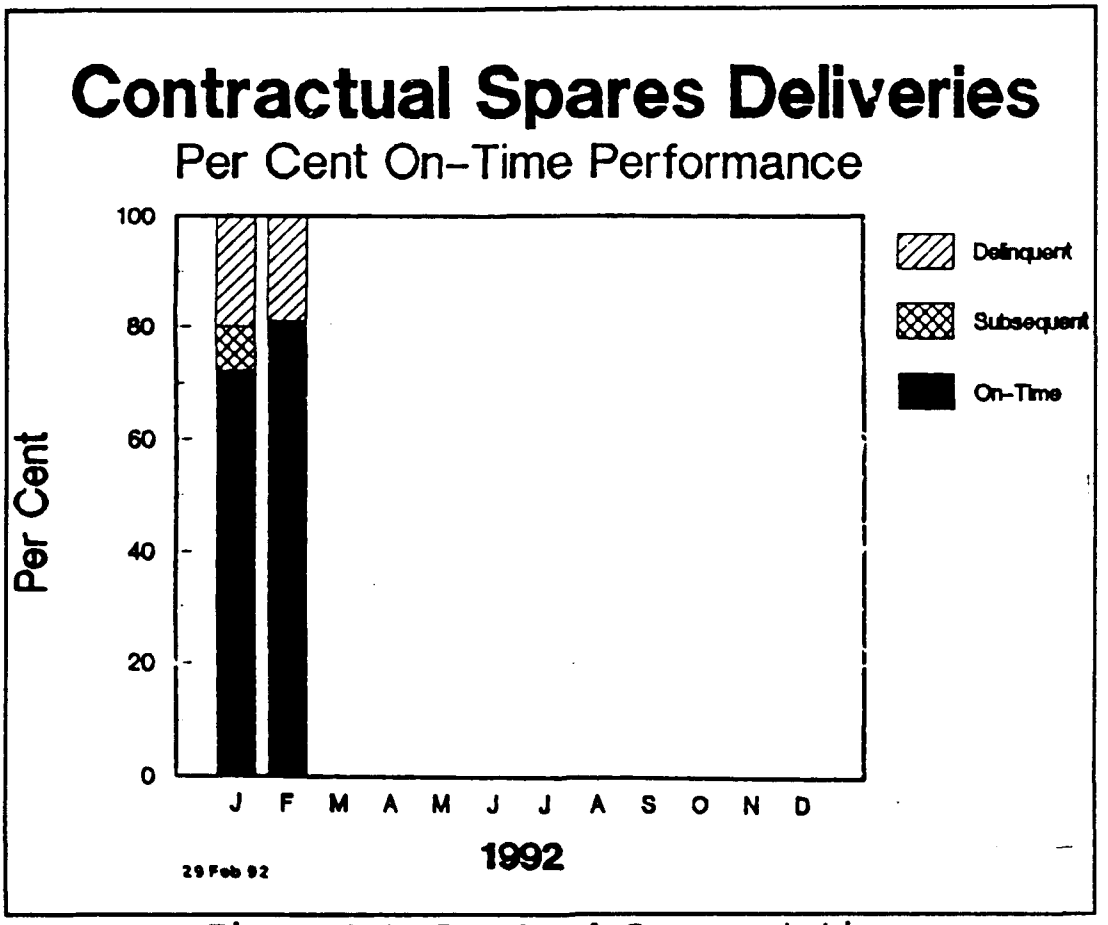


Figure 3-2. Bar Graph Representation

metrics were first organized and then selection criteria were applied to identify the metrics that would be used in the evaluation.

Metric Taxonomy. The research team systematically organized the metrics into the categories of cost, schedule, and performance. These three categories are recognized as being the three most important to an acquisition project. Because of the number of metrics in each of these categories, a single metrics category, schedule, was identified as the focus of the research. Schedule metrics

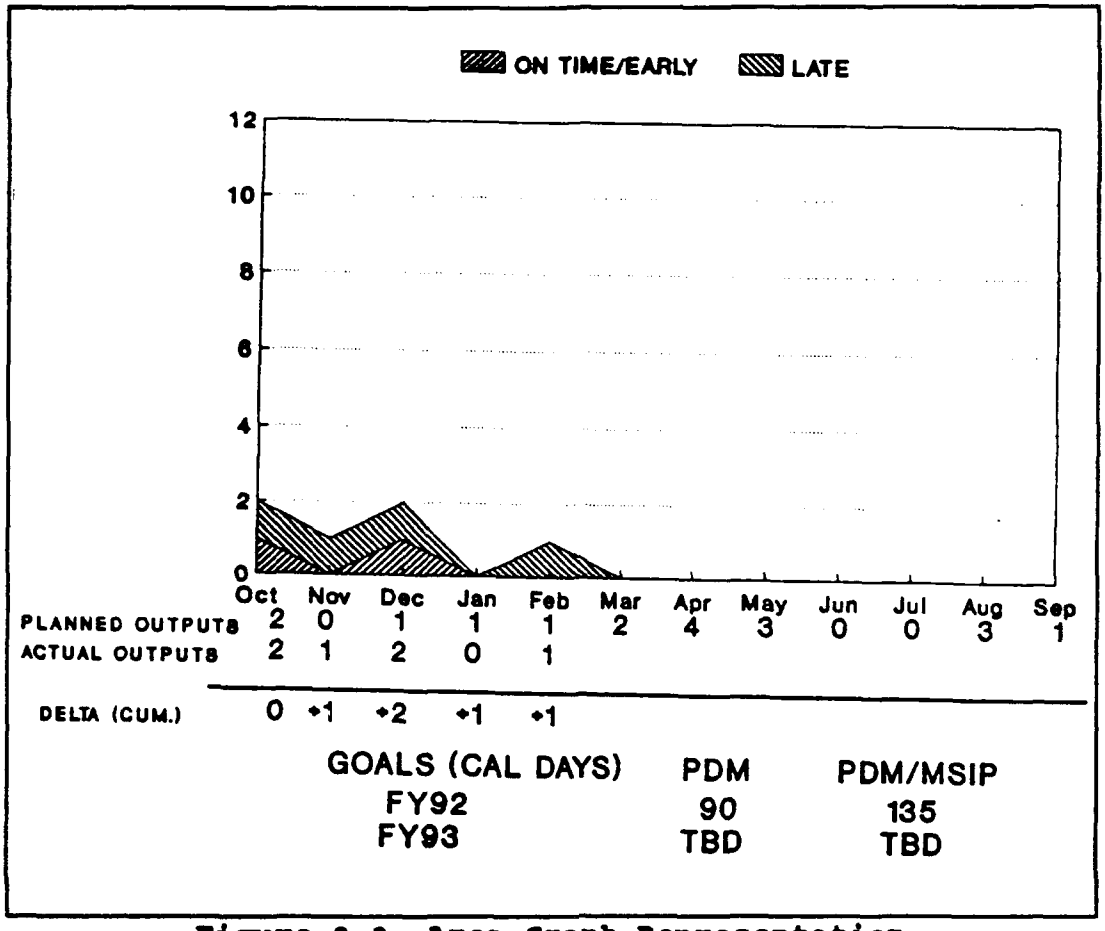


Figure 3-3. Area Graph Representation

were selected because they were more amenable to classification and were those most common to all the SPOs identified for the study. Thus, the schedule metrics provided a more logical structure from which metrics could be chosen for evaluation. Specifically, two categories of schedule metrics were identified: delivery schedules and program schedules.

Metric Selection Criteria. The next step was to choose the actual metrics from these two categories. Four criteria were used to choose the metrics. The first criterion used

in choosing the metrics was their general applicability to many program offices. In other words, generic-type metrics were selected. Secondly, the metrics chosen had to be important to the success of the programs in which they were being used. Insignificant metrics or those that measured processes that did not directly affect program success were not considered. The third criterion was that the metrics chosen must be sufficiently different from each other to ensure coverage of a diversity of SPO activities. Finally, the metrics should be those that contributed to meeting typical SPO objectives, such as meeting baseline objectives.

Selected Metrics. Using the criteria listed above, seven generic schedule metrics were selected for evaluation. Three delivery schedule and four program schedule metrics were chosen to ensure adequate coverage of the types of schedule metrics. The metrics that were selected, along with a brief explanation of each follows.

Metric 1. Number of units of product actually delivered versus the number of product required to be delivered on a monthly basis.

Delivery in this case is to the user. Required means contractually required. Units of product can be aircraft, pods, radios, or any other hardware or software item. This metric results in two numbers used for comparison.

Metric 2. Average number of days to respond to customer service report requests on a monthly basis.

Service reports are problems or issues identified by the user for corrective action. Response in this case would be the contractor or government providing an initial answer to the problem. An average is used because there may be numerous reports generated per time period.

Metric 3. Number of undefinitized contractual actions (UCAs) definitized within 180 days versus those definitized after 180 days.

UCAs are contractual actions which do not have an agreed upon price. Work is allowed to start using a not-to-exceed price in order to limit government cost risk. They are used because it is often desirable to begin the work immediately rather than wait until a formal written agreement is reached. A UCA becomes definitized when subsequent agreement is reached on the exact changes to contract pricing. One hundred-eighty days is a requirement mandated by federal law. This metric results in two numbers: the number of actions definitized within 180 days and the number definitized after 180 days.

Metric 4. Number of project baseline breaches per month.

A breach means that the schedule in the finalized baseline was not met. In this case, baseline refers to an informal agreed upon plan of attack to meet schedule goals (as opposed to the formal baseline for which breaches must be reported to higher headquarters). This metric results in single number.

Metric 5. Average variance in days between the scheduled release of Requests For Proposals (RFP) and the actual RFP releases by month.

A variance is the number of days between the planned release of an RFP to prospective contractors and the actual date the RFP is released to prospective contractors. An average is used because there may be multiple RFPs released per time period.

Metric 6. Number of contractor data submittals due versus those actually received on a monthly basis.

Contractor data submittals can be hardware or software specifications, test reports, operation manuals, cost data, etc. In this case, received refers to receipt by the cognizant office within the SPO. Due means the contractually specified date due. This metric results in two numbers for comparison.

Metric 7. Number of actual Government Furnished Equipment (GFE) deliveries provided versus the number actually required on a monthly basis.

GFE is equipment owned by the government but contractually required to be furnished to the contractor for use during the acquisition process. This metric results in two numbers for comparison.

After the metric selection process, the research team addressed the criteria and process that would be used to evaluate the metrics.

Metric Evaluation

Evaluation Criteria. A literature review was performed to identify the criteria that constitute a good metric as well as illustrate the importance of metrics. Total Quality Management (TQM) philosophies both inside and outside the government were analyzed for metric applicability. The team scrutinized national quality award programs as well as writings of a number of acknowledged quality experts. In addition, current ASC policy and literature were examined to determine the current thinking on the characteristics of a good metric.

A number of criteria for "good" metrics were identified. However, in an effort to narrow the scope of the research, the team focused on two rating criteria -- how well the metric drives a behavior and how well the behavior contributes to continuous improvement. Throughout the literature review process, research clearly showed that driving the right behavior and continuous improvement were the key components of a good metric. These components are the basis for the development of ASC metrics (1:21).

Evaluation Method. The following sections present the process used to construct the evaluation methodology.

Group Decision Making. In order to evaluate the metrics, the research team decided that a group process should be used. Groups tend to yield more accurate judgments than individuals (20:99, 27:318, 19:329).

Nominal Group Technique (NGT) is one method of implementing group decision-making. NGT has been shown to be superior to other group decision making techniques for producing more accurate and better quality judgments (27:319, 20:98).

> NGT is a method for brainstorming in which participants first submit ideas anonymously with no discussion or criticism of anyone's ideas allowed. After ideas are collected, the group then discusses them and selects the best ideas to pursue. (5)

A true NGT process consists of four steps:

> 1. silent judgments by individuals in the presence of the group,
> 2. presentation to the group of all individual judgments without discussion,
> 3. group discussion of each judgment for clarification and evaluation,
> 4. individual reconsideration of judgments and mathematic combination. (27:319)

A variant of NGT was selected for use in this case. First, steps two and three above were essentially combined. After individual development of ideas, open discussion and questions were encouraged. However, no value judgments were made concerning individual ideas. Conversation was limited to clarification and consolidation of duplicate ideas. This step ensured that all members of the group understood the ideas presented. In addition, the wording of the ideas was refined. In step four, the evaluators did not reconsider the original ideas. Instead, they were given the opportunity to rate each idea. In this way, all the ideas generated were considered.

Group Support System. The next decision was how to implement the NGT process. A Group Support System (GSS) was chosen.

> Group Support Systems are computer-based systems that provide a variety of tools to facilitate the meeting process. In part, these systems are electronic implementations of older methods - e.g. Delphi and Nominal Group Technique - that have been used to improve the quality of meetings over the last 30 years...Over 88% of the users in these studies felt the system had improved the quality of the decisions reached. (5)

The GSS used for the research was located on Wright-Patterson AFB, in the Armstrong Lab. Armstrong Lab primarily performs research into human capabilities and how these effect weapon systems and operational performance (5). Specifically, the lab focuses on "technology for improving performance of integrated systems of people, information, and equipment doing essential acquisition and combat support functions" (5).

The software system used was GroupSystems V produced by Ventana Corporation. GroupSystems is a software tool used to facilitate decision-making. Not only does the software implement NGT in a more effective and efficient manner than non-electronic means, it also provides a report of all decisions including comments from all the group participants. Further, it also allows the group to evaluate various options and produces a numerical summary of results (32). One of the greatest strengths of GroupSystems is its flexibility in adapting to the needs of the meeting.

Screens can be customized to support the topics being discussed.

Group Selection. Having chosen the method of evaluation, the next step was to select the individuals that would perform the evaluation. The primary consideration was that the evaluation be performed by experienced acquisition program managers. Qualifications for the group members included two criteria. First, group members should qualify as a Level II acquisition professional in the Acquisition Professional Development Program (APDP). To qualify for level II, program managers must have completed Systems 200, a Professional Continuing Education (PCE) course provided by the Air Force Institute of Technology (AFIT), and served at least four years in an acquisition related position (24). This criterion focused on the individual's experience as a program manager in the acquisition arena. Because the internal SPO metrics collected focused on day-to-day program management, the research team decided that an evaluation group of "hands on" program managers was most appropriate for the research. Secondly, because continuous improvement is such a key component of Total Quality Management, TQM experience was also considered a key attribute in the group selection process. Therefore, each participant was required to have received some formal TQM training.

Determining the proper size of the group was constrained by the number of computer stations available, in this case ten. However, the most effective size for group

decision making has been found to be something less than ten. Several studies have shown that as group size increases, communication becomes more difficult and some members of the group may tend to withdraw. On the other hand, if a group is too small, participants may be forced to contribute in a manner in which they are not comfortable. As a result, group sizes of five have been found to be very effective for small discussion groups involved in making decisions (19:231, 18:142, 20:86).

Among the program manager population at Wright-Patterson, over 600 program managers met both criteria. The five individuals that participated in the evaluation were chosen by convenience. Four were former coworkers of one of the research team members and the fifth was a current classmate of both researchers.

Group Evaluation Process Development. After selecting the method of evaluation and the participants, the research team refined the evaluation process through participation in a software orientation session and two dry runs prior to the actual evaluation.

Software Orientation. The objective of the orientation was to determine the feasibility of using the software at Armstrong Lab to facilitate evaluation of the metrics. The research team participated in a demonstration of GroupSystems V, the package being used. The software demonstrated the capability to capture individual ideas while simultaneously viewing other's ideas. In addition,

its ability to subsequently edit and organize these ideas, and then rate each of them numerically was demonstrated. The orientation resulted in an understanding of the software and provided an initial concept of the details of the evaluation process.

Development Dry Run. Next, the research team provided the metrics and evaluation criteria to Armstrong Lab so that candidate computer screens for the evaluation could be constructed. After the software had been customized with this information, the first dry run took place. The objective of this was to determine the detailed process that would be used during the evaluation. To facilitate this, one of the metrics was actually evaluated. The research team evaluated each screen in depth to ensure that the text and instructions were clear. In addition, the order of the steps and length of each were examined. This step resulted in the development of the five-step evaluation process that was used.

Validation Dry Run. The second dry run used the results and subsequent refinements of the first dry run and further defined the process. The screens that the participants would use were updated. In addition, the first dry run emphasized the importance of limiting the amount of time spent on each step. This was necessary to ensure that an adequate number of metrics could be evaluated in a single-day session. Therefore, the research team participated in the evaluation of a sample metric. Each

step was performed exactly as it would be during the actual evaluation, including timing each step. The result of this dry run was a finalized process, including the time allotted for each step, that would be used for the actual evaluation. Screen content was refined and the roles of all participants, including the facilitator and software session coordinator, were finalized. The five-step evaluation process that resulted from this dry run is presented in the next section.

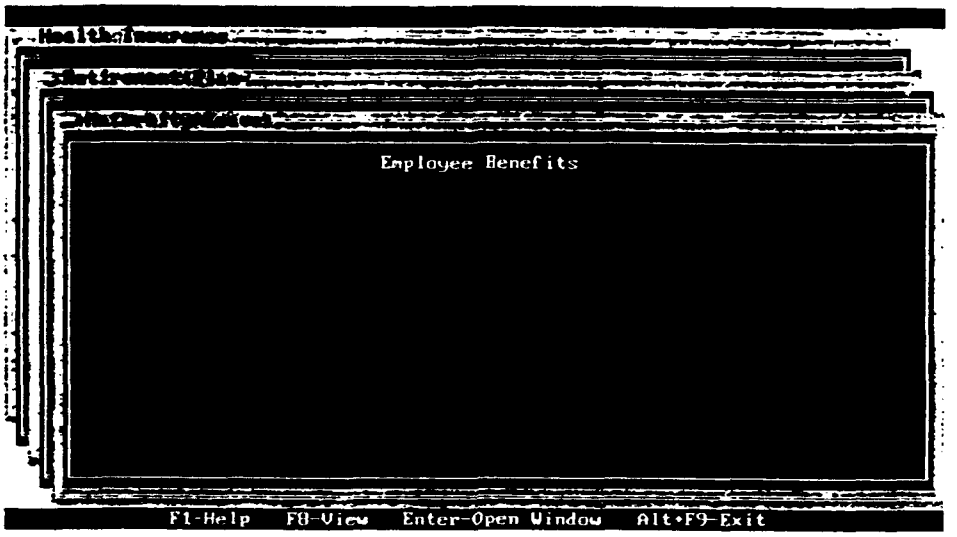
Group Evaluation Process Execution. The evaluation of the metrics was conducted as follows. The group met one day from 0800 to 1600 to evaluate the metrics. The intent of the team was to evaluate approximately one metric each hour. After providing job aids, and refreshments the session began. The evaluation consisted of a five step process preceded by a software orientation session.

Software Orientation. During the orientation, a facilitator familiarized the group with the software they would be using and led them through an example metric. In addition, the facilitator described job aids which were provided to the participants (see Appendix B). The example metric was an actual metric collected from a SPO in ASC and was selected to demonstrate the five-step process the group would use during the metric evaluation. The orientation portion lasted 45 minutes.

Predict Behaviors. The purpose of the first step was to elicit individual opinions of the possible behaviors

that may result from the metric. The facilitator advised the group that they were to comment on the possible behaviors that could be driven by the metrics. No discussions would take place during this first step. In addition, they were told that their responses should be limited to one line, if feasible, to facilitate follow-on discussion and evaluation. In addition, no value judgements were to be included in the behavioral comments. This was required to ensure that all possible behaviors were identified, both positive and negative. If the behavior identified was a specific action, the actor would also need to be identified. Each group member had access to his own dedicated computer terminal. During this step of the evaluation process, the group entered one line responses about the possible behaviors the particular metric could drive. Figure 3-4 is an example of the Topic Commenter screen in which these comments were entered. The behavioral comments were also sent anonymously to a large wall screen in the front of the room which could be viewed by all. In addition, they had the opportunity to observe other group member's comments (again, anonymously) on their own computer screen. This provided an additional opportunity to provide comments that may have been prompted by another group member's earlier comment. This first step in the session was limited to 10 to 15 minutes. The outcome of step one was an unedited list of possible behaviors that might be driven by the metric.

An example of a participant screen appears as follows:



Entering Comments

The heart of this tool is the editing screen on which participants enter their comments about a topic. Once comments are entered, you can edit them. (See *Editing Comments* later in this chapter for more information.)

❒ *DIRECTIONS TO PARTICIPANTS:*

☞ Note: Topic titles can be 72 characters in length, but only 53 characters will be displayed on your screen. If the topic title is longer than 53 characters, an arrow on the right and/or left border of the title indicates that there is additional text. Use the left and right navigation keys to scroll.

1. **Using the up and down navigation keys, highlight the topic card upon which you want to comment and press <Enter>.**

Figure 3-4. Topic Commenter Screen

Refine the List of Behaviors. The purpose of step two was to clarify and refine the list of behaviors developed in step one. In this step of the process, the facilitator encouraged group discussion. He asked the group to comment on any behaviors they didn't understand. Participants had the opportunity to explain their comments in order to foster a better understanding of their behavioral comments. Any behaviors which were not understood by the group were either clarified or deleted (by mutual agreement). In addition, redundant behaviors were eliminated. This step was also used to ensure that the comments were, in fact, behaviors and not outcomes and that no value judgments were included. Maximum time for this step was 20 minutes. The outcome of this step was a list of clearly understood behaviors that were collectively exhaustive and mutually exclusive.

Rate the Behaviors. The purpose of the third step was to evaluate the list of behaviors that resulted from step two. The facilitator asked the group to rate the behaviors and the metric according to the following questions: 1. How well does the metric drive the identified behavior? and 2. How well does the behavior contribute to continuous improvement? These two questions were rated for each of the behaviors. The rating scale was part of the software package and provided a 10 point scale ranging from very poorly (1) to very well (10). An example of the evaluation screen is shown in Figure 3-5. Maximum time for

Rating Alternatives

❒ *DIRECTIONS TO PARTICIPANTS:*

The criterion window is the "anchor" window. This means that the alternatives window will scroll through all alternatives for the selected criterion. The anchor window is grey in color and has a white arrow to the left of it. You can use the left and right arrow keys to scroll through the list to select another criterion.

☞ Note: If you are reentering the session, the Ratings Summary screen appears. If you previously exited without sending ratings, the following message appears:

```
These ratings have not been sent.

Press any key to continue...
```

1. <Enter> to display the rating scale window.

A rating scale window appears at the bottom of the screen:

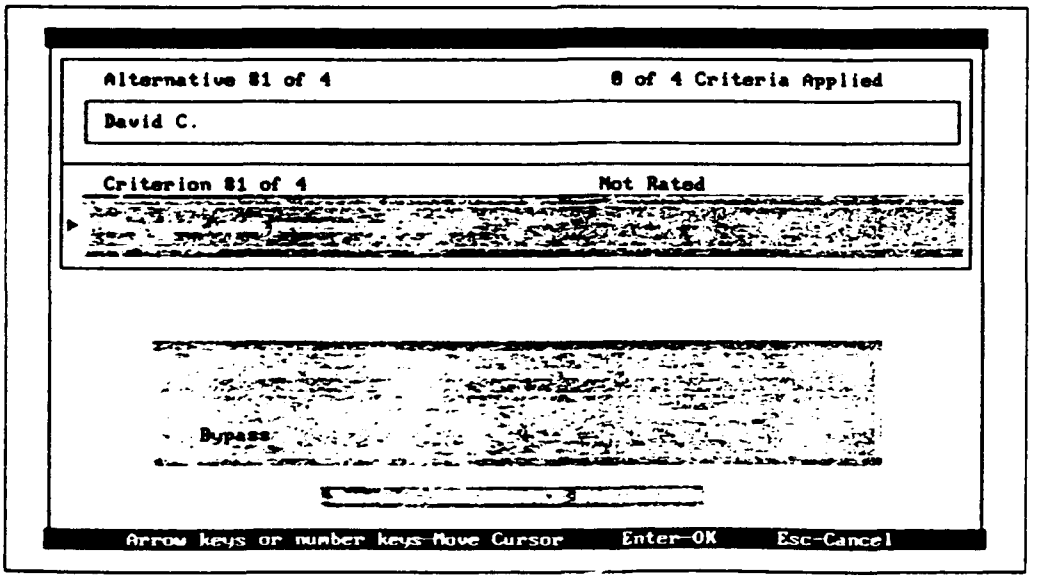


Figure 3-5. Alternative Evaluator Screen

this step was five minutes. The outcome of this step was a numerical rating of both criteria for each identified behavior.

Review the Ratings. The purpose of this step was to allow the participants to see the cumulative results of the group's ratings for each metric. During this step, participants observed the final group ratings via a "cloud" chart presented on the front screen. This chart was presented as a four quadrant graph and represented the group consensus concerning a predicted behavior. Figure 3-6 is an example of a cloud chart showing the results for one behavior. In addition, each participant could observe his individual rating and compare it with the other group members. If desired, the group members could explain their ratings at this time. However, the ratings for that behavior were not subject to change. The group reviewed one cloud chart for each predicted behavior identified for the metric being considered. The outcome of step four was an understanding of how the group evaluated each behavior and clarification of reasons for any major differences in the ratings.

Provide Final Comments. The purpose of the last step was to provide an opportunity for the participants to make any additional comments about the metric or behaviors that may not have been captured earlier. In addition, comments about the process were solicited. The facilitator encouraged the evaluation group to expand on any ideas the

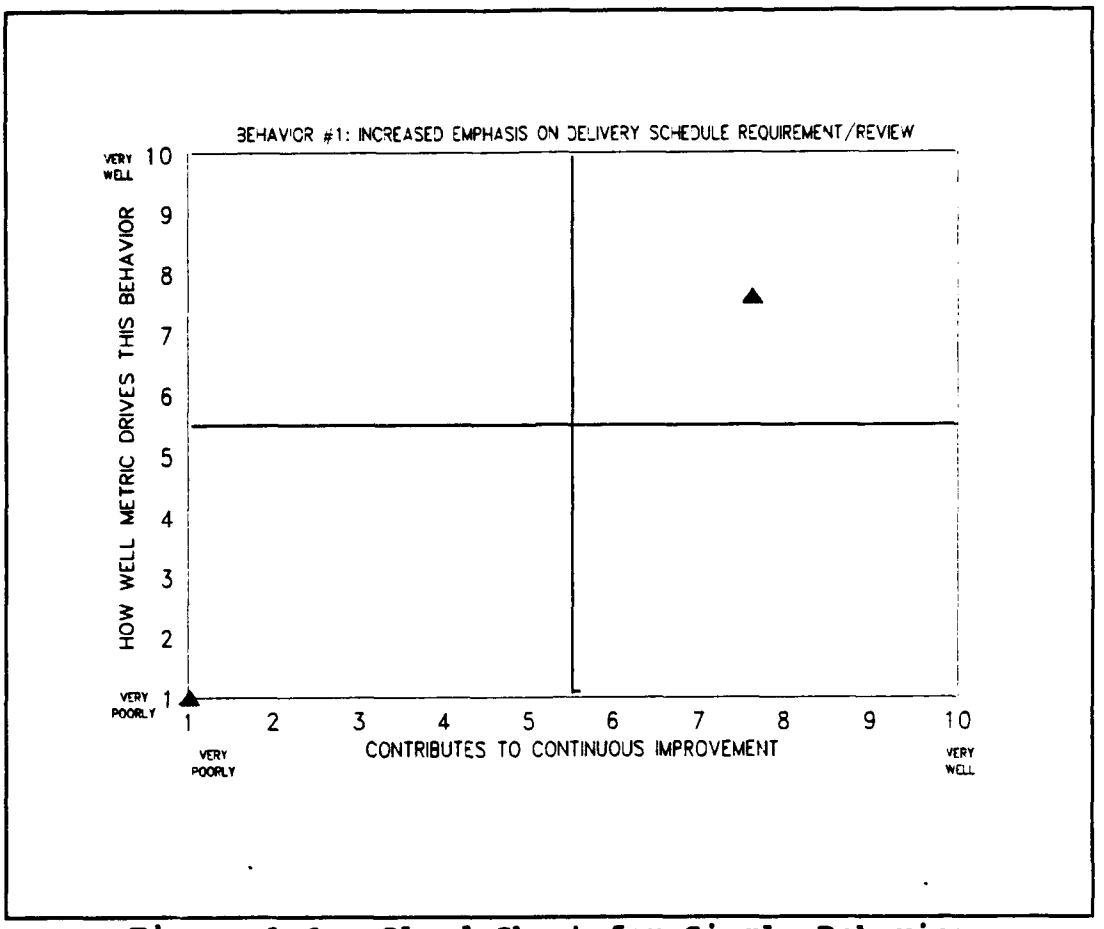


Figure 3-6. Cloud Chart for Single Behavior

first four steps prompted. Specifically, they were asked to comment on the questions previously presented to them: What are the likely behaviors that result from this metric?, How well does the metric drive this behavior?, and How well does the behavior contribute to continuous improvement? In addition, they were given the opportunity to respond in an "ad hoc" manner to any part of the metric evaluation process they experienced. Any comments concerning the metric, its development, or its use were also encouraged. The result of this last step of the process was a list of comments about

the metric itself, its strengths and shortcomings, observations about specific behaviors, and ideas on how to improve the metric. In addition, comments about the evaluation process itself were captured.

Results of the Evaluation

The results of this five-step process were used to analyze each metric. The ratings associated with each of the behaviors can be summarized on a single graph which can reveal an interesting summary of the metric. Each behavior is identified by one triangle on the chart. The triangle is the plot of the mean of the group responses to each question. For example, the triangle in Figure 3-6 represents a mean response of 7.8 to the question "How well does the metric drive this behavior?", and a mean response of 7.6 to the question "How well does this behavior contribute to continuous improvement?". The graph is divided into four equal areas called quadrants. Figure 3-7 illustrates the four quadrants and the case where the behaviors fall in quadrant one. A good metric would have most of the behaviors in quadrant one of the graph. This is because the behaviors are being strongly driven by the metric and are contributing to continuous improvement as well. An ideal metric is one in which all the behaviors fall in quadrant one. Conversely, a poor metric would have many behaviors in quadrant two. A majority of the behaviors in the upper left quadrant would reveal that the metric is

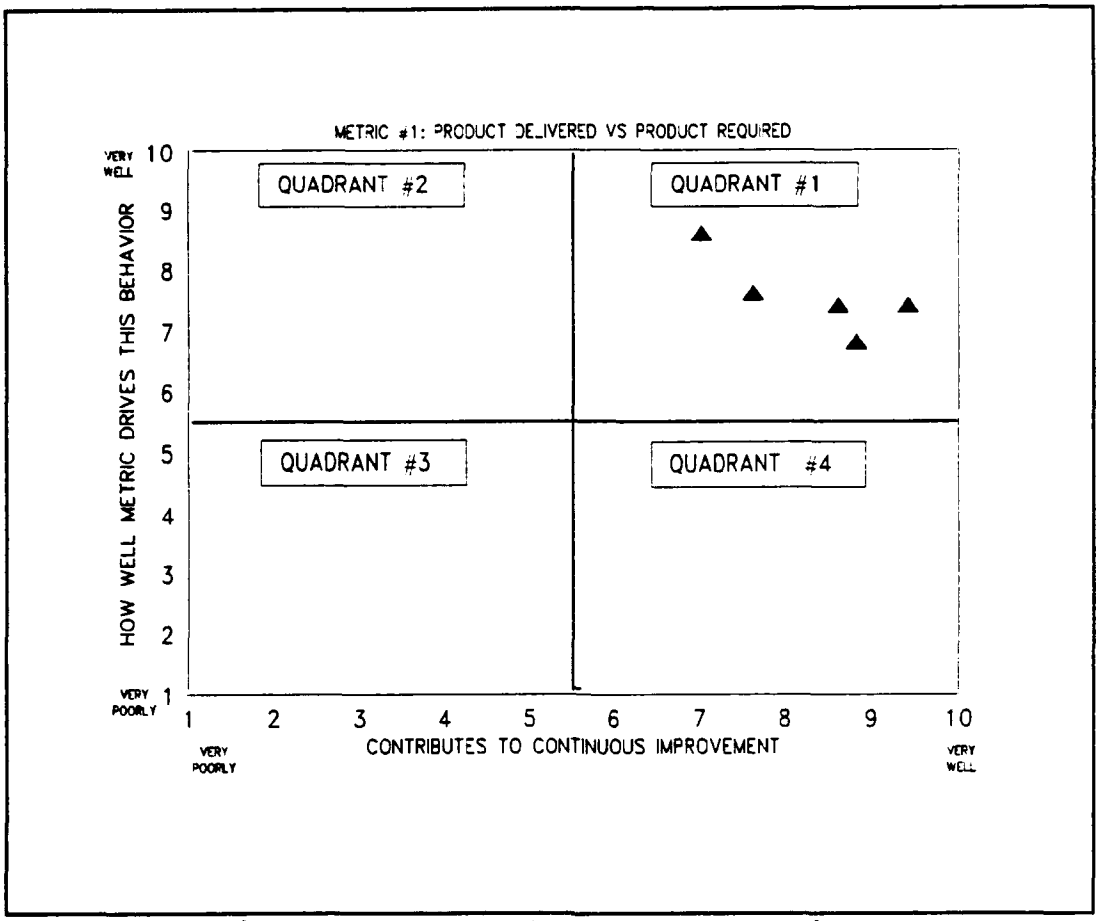


Figure 3-7. Quadrant One Behaviors

driving behaviors that do not contribute to continuous improvement. These are undesirable metrics because they are driving the wrong behaviors. Figure 3-8 illustrates this case. A preponderance of behaviors that fall into quadrant three of the graph indicate an inconsequential metric. In other words, the metric is driving few, if any, behaviors strongly and the behaviors are not contributing to continuous improvement. This is an example of a metric that is clearly ineffective and should not be used. Figure 3-9

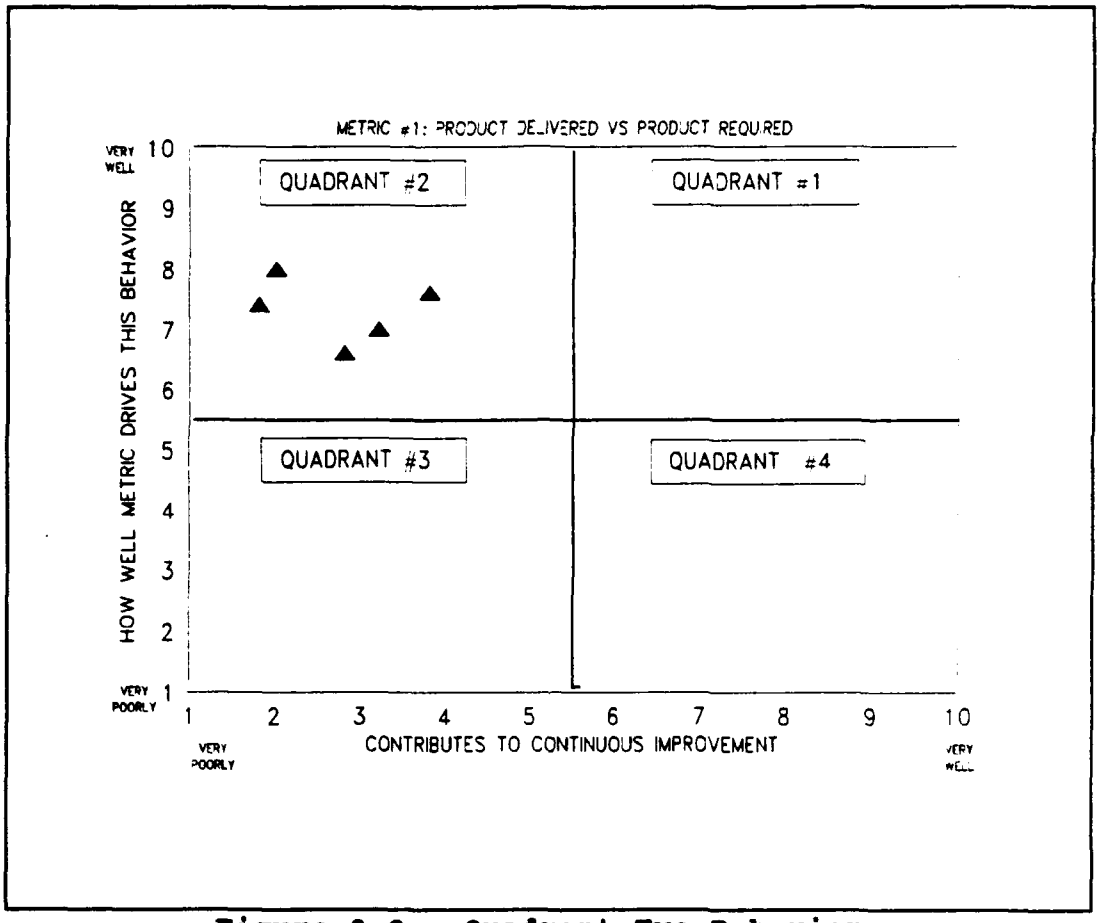


Figure 3-8. Quadrant Two Behaviors

illustrates this case. Finally, many data points in quadrant four indicate behaviors that are desirable but are not being driven by the metric. This is also an example of an inconsequential metric. However, this information can be useful to the developers and users of metrics. These behaviors could be the basis for building a new metric that would, in fact drive the behaviors necessary for continuous improvement. Figure 3-10 illustrates this case.

The research team did not expect to encounter any metrics that fit neatly into a single quadrant. The

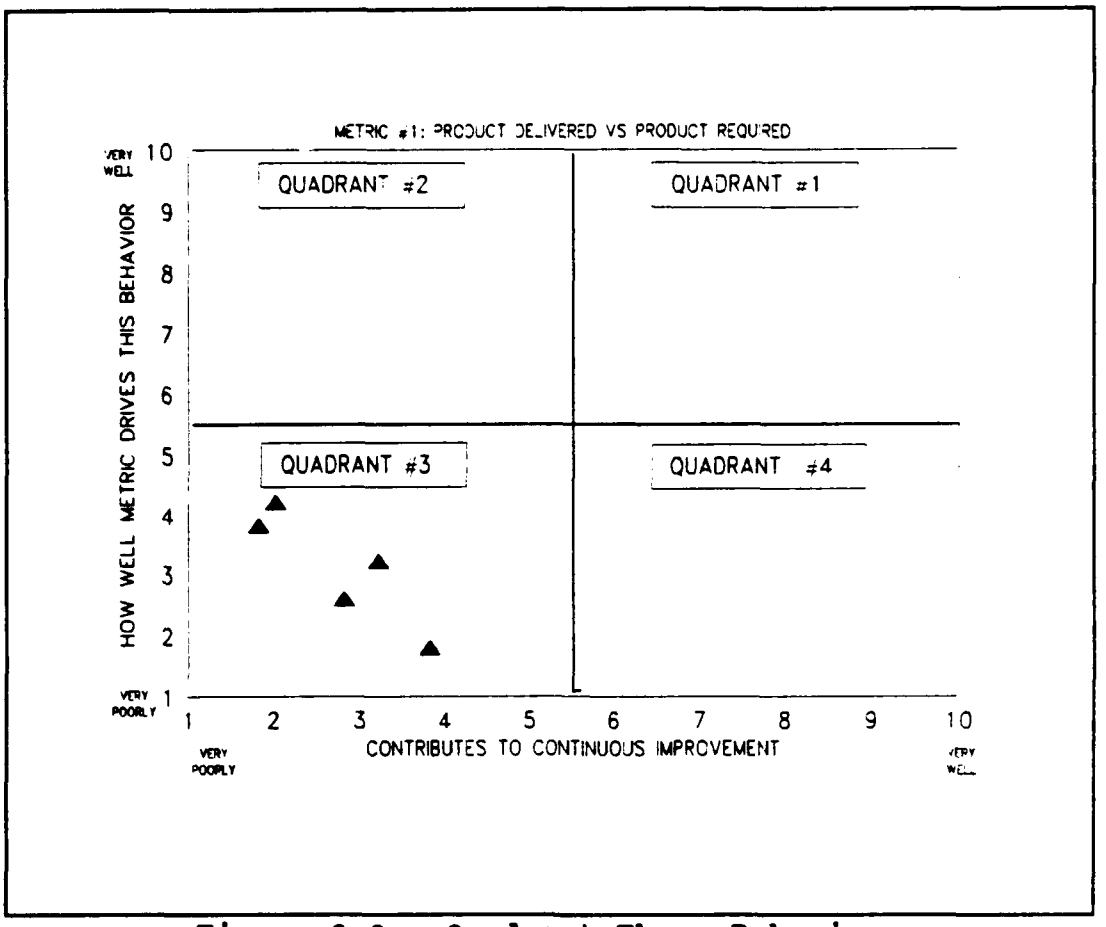


Figure 3-9. Quadrant Three Behaviors

expectation was that the behaviors would be scattered throughout the quadrants. This is for two reasons. First, the development of internal-SPO metrics is still in it's infancy. Therefore, behaviors that are not linked to continuous improvement may be driven by the metrics due to lack of experience. Second, the evaluators of the metrics have different levels of acquisition experience and knowledge of metrics. This will result in different interpretations of the same metric, or even the same behavior.

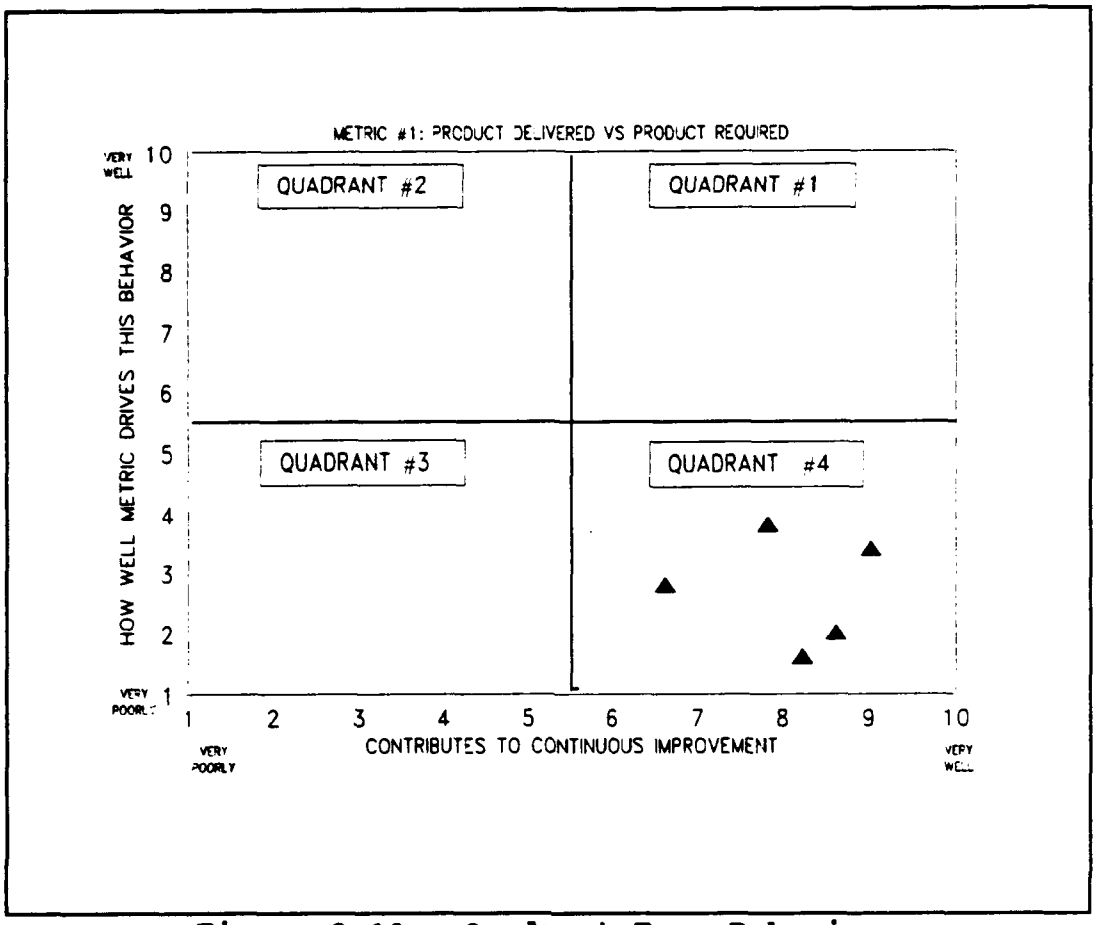


Figure 3-10. Quadrant Four Behaviors

In addition, the research team expected that most of the behaviors would fall into the upper quadrants, one and two. This is due to the nature of the design of the evaluation process. Steps one and two were designed to identify likely behaviors. Therefore, most of the behaviors identified should, in fact, have a high rating on the question of how well the metric drives the behavior. This would result in most behaviors falling into the upper quadrants of the graph.

Using the results of the process described above, the behaviors for each metric were combined to form a single cloud chart. This is the basis for the analysis of the individual metrics that is presented descriptively in the next chapter.

Summary

Table 3-1 illustrates the mapping of the methodology to the research questions posed in chapter one.

TABLE 3-1

RESEARCH METHODOLOGY SUMMARY

RESEARCH QUESTION	METHODOLOGY	PAGE #
What are the criteria of a good metric?	Literature Review	3-12
What are the internal schedule metrics currently in use within ASC SPOs?	Mail Survey	3-2
What are the possible behaviors that are driven by these schedule metrics?	Small group evaluation using a Group Support System	3-13
Do the behaviors driven by the schedule metrics lead to continuous improvement?	Small group evaluation using a Group Support System	3-13

IV. Metrics Evaluation and Analysis

Using the selection process explained in the previous chapter, seven generic metrics were chosen for evaluation. The metrics are listed in Table 4-1.

TABLE 4-1

LIST OF METRICS

1. Number of units of product actually delivered vs. the number of product required to be delivered on a monthly basis
2. Average number of days to respond to service report requests on a monthly basis
3. Number of undefinitized contractual actions (UCAs) definitized within 180 days vs. those definitized after 180 days
4. Number of project baseline breaches per month
5. Average variance in days between the scheduled release of Request for Proposals (RFPs) and the actual RFP releases by month
6. Number of contractor data submittals due vs. those actually received on a monthly basis
7. Number of Government Furnished Equipment deliveries provided vs. the number actually required on a monthly basis

Metric 1

The first metric evaluated was "Number of units of product actually delivered vs. the number of product required to be delivered on a monthly basis. First, the evaluation group was asked to answer the following question:

What are the likely behaviors that result from this metric? Table 4-2 contains the list of behaviors identified by the evaluation group.

TABLE 4-2

LIKELY BEHAVIORS RESULTING FROM METRIC 1
NUMBER OF PRODUCT DELIVERED VS. NUMBER OF PRODUCT REQUIRED

1. Increased emphasis on delivery schedule requirement/reviews
2. People will increase deliveries near the end of the month
3. Increased emphasis on project planning
4. People will seek a high quantity requirement if they're good workers
5. People will change the schedule deliveries to meet output
6. Increased emphasis on critical supplier schedules
7. People will increase the amount of overtime to meet required deliveries
8. Increased emphasis on critical assembly processes (schedule sensitive)
9. People will increase finger pointing
10. People will seek to ensure delivery process is reliable
11. People will concentrate on delivering exact number required
12. People will become more concerned with quantity than quality
13. Increased emphasis on component availability during assembly process
14. People will pressure on customers to accept product
15. Increased emphasis on statistical process control

Rating Summary for Metric 1. The combined rating for the first metric is shown in the composite graph in Figure 4-1. The individual graphs are located in Appendix A.

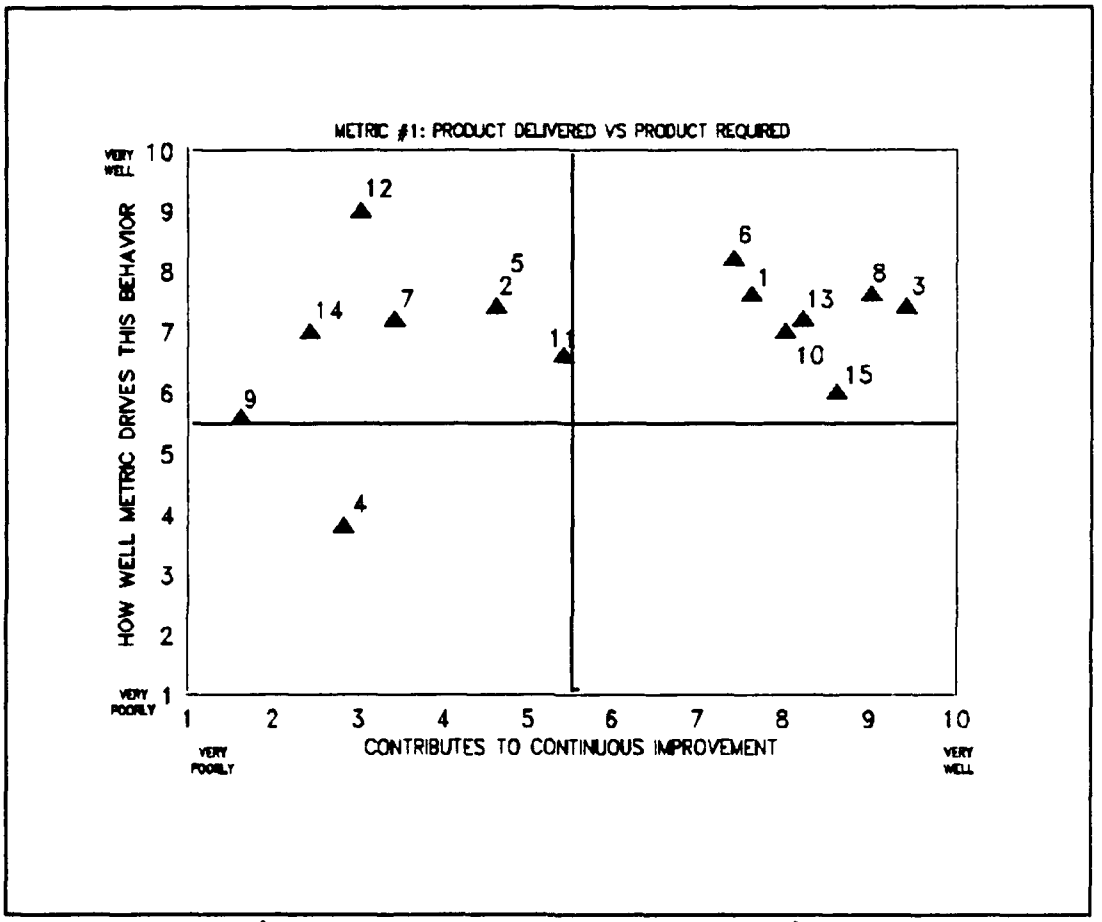


Figure 4-1. Results of Metric 1

Evaluation Group Comments. The evaluation group recorded the following comments about the first metric following the rating of the behaviors.

> This metric was too open-ended. The metrics need to be more specific in nature to assure that one is in fact measuring correct behaviors.

In order to better elicit the proper behaviors, other metrics need to be used in conjunction with this one. It will add focus to the metric.

Metric formation should include those being measured in the process.

It is difficult to judge this metric without other metrics which could be used simultaneously. There is a tendency to only see the misuse of this metric, which could be easily corrected by adding a second metric. For example, to prevent quality from suffering, include a metric to measure scrap and re-work, customer satisfaction, etc.

This metric is an important part of SPO management activities. The metric is a bit too general and non-specific to gather information useful to assess the specific behaviors you may be looking for.

Even though this process has identified some very good behaviors for this metric, a little more work on the metric would produce possibly better results/behaviors.

The discussion on whether the behavior should apply to contractor/SPO or generally to customer/suppliers (especially internal) reflects the typical mentality which puts the blame on contractors or others rather than holding ourselves and our internal processes accountable. People need to recognize that products and services are continually being delivered between two people no matter how big or important or "real" the product.

Analysis. This metric exhibits two distinct characteristics. From Figure 4-1, it is apparent that this metric is driving most of the likely behaviors in a moderate to strong manner. However, many of the likely behaviors are not associated with continuous improvement. Eight of the fifteen behaviors fall in the quadrant one. This indicates that the metric is both appropriate and consequential. This seems to be particularly true for those behaviors that are

driving people to examine the process itself, rather than the output required. For example, for those behaviors that mention increasing emphasis on aspects of the process that are critical to meeting delivery requirements, ratings were especially high for both criteria. Behavior three, increased emphasis on project planning, focuses on improving the planning process that leads to timely delivery of products. Rather than looking at the result of the process, the process itself is being examined.

On the other hand, behaviors that focused on the end product exclusively fell into quadrant two, indicating behaviors likely to be driven but which will not contribute to continuous improvement to a large degree. These behaviors are those that often focus on meeting delivery goals and quotas. For example, behavior fourteen, putting pressure on the customer to accept the product, concentrates on the product itself. No mention is made of whether or not the product meets quality requirements and is acceptable to the customer. Behaviors seven and twelve focus exclusively on meeting quantity requirements. Again, quality and customer satisfaction are not considered. According to Deming, a focus on goals is useless for two reasons. First, if the process is stable to begin with, maximum quantities are already being achieved. Attempting to force any additional production will only introduce variation to the process, resulting in decreased rather than increased production. Secondly, if the process is not stable to begin

with, attempting to increase production will simply introduce more variation into an already unstable process, resulting in a deteriorated process. The process must be improved before production improvements can be achieved. This must be accomplished by management, not the production workers. The behaviors that fall into quadrant two address functions of workers at the production level. Conversely, the behaviors in quadrant one address functions that are normally in the management realm. These are the behaviors that can change the process, enabling production workers to meet requirements without introducing destabilizing quotas and variation into the process (11:65-69).

It is also apparent from the evaluation group's comments that this metric is difficult to completely evaluate unless it is used in conjunction with other metrics. Therefore, some of the behaviors that are being driven, but not necessarily in the direction of continuous improvement, may be desirable when combined with behaviors from other metrics. For example, if a metric that measured the quality of the product was used with this metric, behaviors such as people being more concerned with quantity than quality would be eliminated.

Another difficulty in evaluating this metric was it's generic nature. A more detailed explanation of the metric might lead to a more accurate list of behaviors that are being driven. Specifically, without knowing the objective to which this metric is linked, it is difficult to make

completely accurate judgments. Additionally, it is important to know who is performing the behavior being driven. Is it the contractor? Is it the SPO? The behaviors being driven depend on who thinks they are being measured by the metric. If people believe they are being evaluated by this metric but actually are not, the metric may drive behaviors that are not in line with the objective of the metric. This is another reason why precise definition is critical.

Metric 2

The second metric evaluated was "Average number of days to respond to service report requests on a monthly basis". Table 4-3 contains the list of likely behaviors this metric may drive.

TABLE 4-3

LIKELY BEHAVIORS RESULTING FROM METRIC 2
AVERAGE NUMBER OF DAYS TO RESPOND TO SERVICE REPORT

1. Contractor will develop a process to reduce response time
2. Contractor emphasis on response time instead of response content
3. Contractor shifts manpower to quickly work the SRs
4. Increased SPO service report status reviews
5. SPO will pressure contractor to respond faster
6. SPO will develop a process to resolve a higher rate of SRs
7. SPO emphasis on contractor response time instead of SPO SR resolution

TABLE 4-3 (CONT)

LIKELY BEHAVIORS RESULTING FROM METRIC 2
AVERAGE NUMBER OF DAYS TO RESPOND TO SERVICE REPORT

8. Contractor tries to get heads-up before problems are officially submitted
9. Customer will feel his SRs are important; may increase SRs

Rating Summary for Metric 2. The combined rating for the second metric is shown in the composite graph in Figure 4-2. The individual behavior graphs are located in Appendix A.

Evaluation Group Comments. The evaluation group recorded the following comments about the second metric following the rating of the behaviors.

> Although response time may be reduced, the quality of the review, i.e. concentration of response content, may not decrease.
>
> This metric is really misleading. It is good to be responsive. However, what is not measured is resolution of system problems that caused the service report to be generated.
>
> This metric seems to be encouraging fire-fighting and suboptimization.
>
> If this is the only metric used, people's focus will not be on what's really important, i.e. long term problem resolution and development of a product that doesn't require service reports to be generated.
>
> This metric does not include "and resolve" which is a CRITICAL part of the SR process. My experience indicates this is by far the most

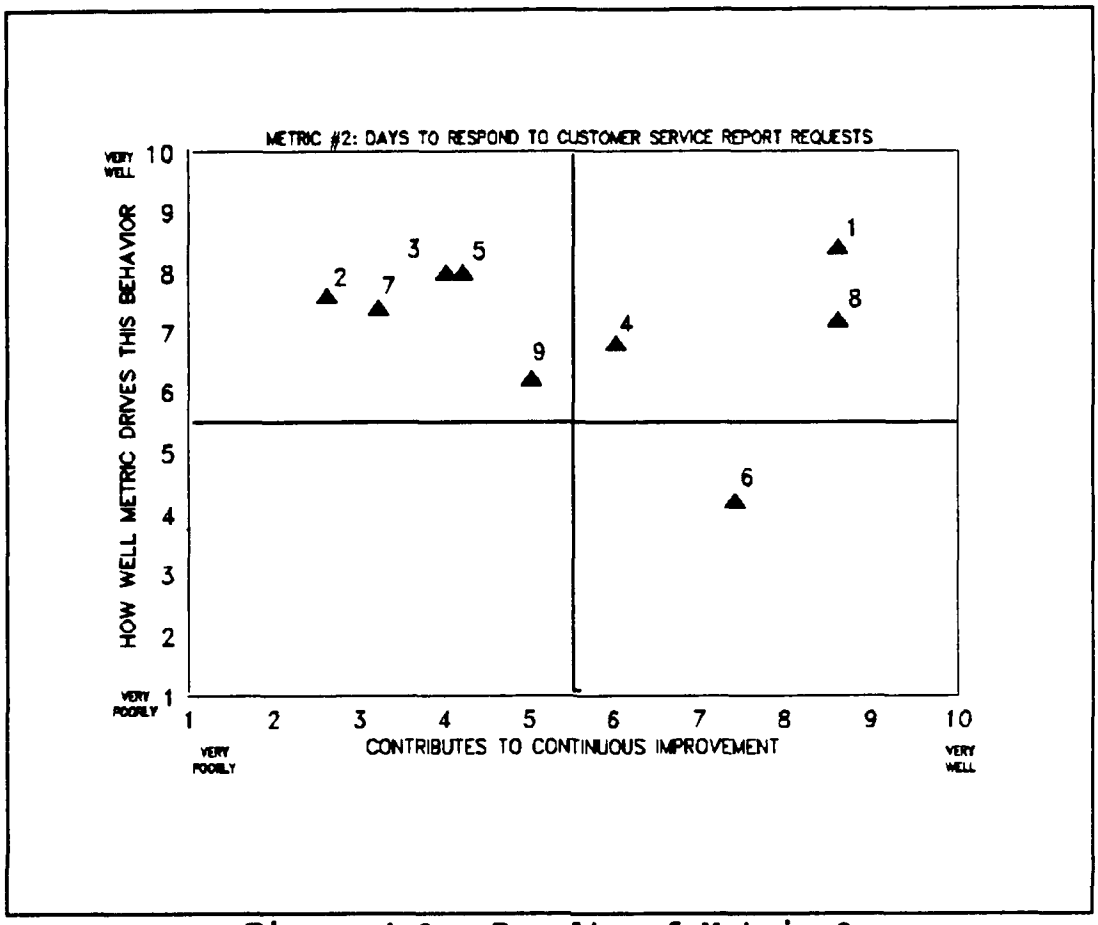


Figure 4-2. Results of Metric 2

troublesome portion of the process and that an "SR Resolution" metric would be good to use.

Another point which came out in the discussion was an indication that the SPO may not consider SRs to be all that important. No wonder the testers and customers frequently don't write SRs; they suspect we just file them and don't do anything about their concerns anyway.

The contractor cannot work in a vacuum on decreasing response time.

The government must increase the quality of the SR in order to facilitate the contractors response. In other words, be specific with the description of the problem. A metric to elicit this behavior would benefit the response time metric.

Analysis. The results of the evaluation indicate there is a rough split in the behaviors that are contributing to continuous improvement and those that are being driven but do not contribute to continuous improvement. While this metric is clearly driving some desirable behaviors, it is also motivating some undesirable behaviors. For example, behavior one, developing a process to reduce response time, is clearly focussing on improving the response time for the customer. On the other hand, behaviors two, three, five, and seven, concentrate on getting a response as quickly as possible, with little regard for the quality of the response. Worse yet, the satisfaction of the customer is overlooked. This dichotomy may result because the metric emphasizes measures of response rather than resolution. In this case, the quality of responses may suffer due to excessive emphasis on providing a timely response, regardless of content. A more appropriate metric might contain a measure of the quality of response in addition to the speed in which it is provided.

An even better metric in this case might focus on the resolution of the service report. Measuring the time it takes to receive a satisfactory response from the customer's point of view would probably be more worthwhile. If a measure of customer satisfaction was used, it would be critical to identify the customer. In its current form, the SPO is the customer and the contractor is held responsible for the response. It would be more effective if the metric

measured the time it took to resolve the problem from the date the user submitted the service report.

Metric 3

The third metric evaluated was "Number of undefinitized contractual actions (UCAs) definitized within 180 days vs. those definitized after 180 days". Table 4-4 contains the list of likely behaviors this metric may drive.

TABLE 4-4

LIKELY BEHAVIORS RESULTING FROM METRIC 3
NUMBER OF UCAS DEFINITIZED WITHIN 180 DAYS
VS. NUMBER DEFINITIZED AFTER 180 DAYS

1. Increased SPO contracting reviews for UCA activities
2. SPO will decrease detail of technical evaluation
3. Increased emphasis on SPO UCA definitization process
4. SPO emphasis on striking ANY deal when approaching the 180 day limit
5. Lots of activity as the 180 day limit approaches
6. SPO negotiates faster with less concern for price when approaching 180 day limit
7. Contractor will slow down negotiating process to corner government
8. Increased emphasis on contractor UCA proposal development process
9. SPO UCA definitization priority established by UCA age
10. SPO will decrease use of UCAs unless hard core requirement really exists
11. Decreased proposal evaluation periods
12. Increased manpower working UCAs older than 180 days

TABLE 4-4 (CONT)

LIKELY BEHAVIORS RESULTING FROM METRIC 3
NUMBER OF UCAS DEFINITIZED WITHIN 180 DAYS
VS. NUMBER DEFINITIZED AFTER 180 DAYS

13. Quicker negotiations
14. Other program activities placed on hold to finish definitization actions
15. Contractor/SPO partners perform more up-front agreement on required actions

Rating Summary for Metric 3. The combined rating for the third metric is illustrated in the composite graph in Figure 4-3. The individual behavior graphs are located in Appendix A.

Evaluation Group Comments. The evaluation group recorded the following comments about the third metric following the rating of the behaviors.

> It is impossible to rate "How well does the behavior contribute to continuous improvement?" without knowing or assuming the objective of the metric. In each case, we assumed an objective which, as can be seen by the data scatter, varied from rater to rater.
>
> It would probably be helpful to provide objectives for each metric to get a better understanding of the responses received in the "continuous improvement" evaluation criterion.
>
> There are differing opinions, not even in the same direction, as to what constitutes continuous improvement. This is not the same as having the objective being sought. Furthermore TQ is not just continuous improvement - especially continuous improvement without staying close to

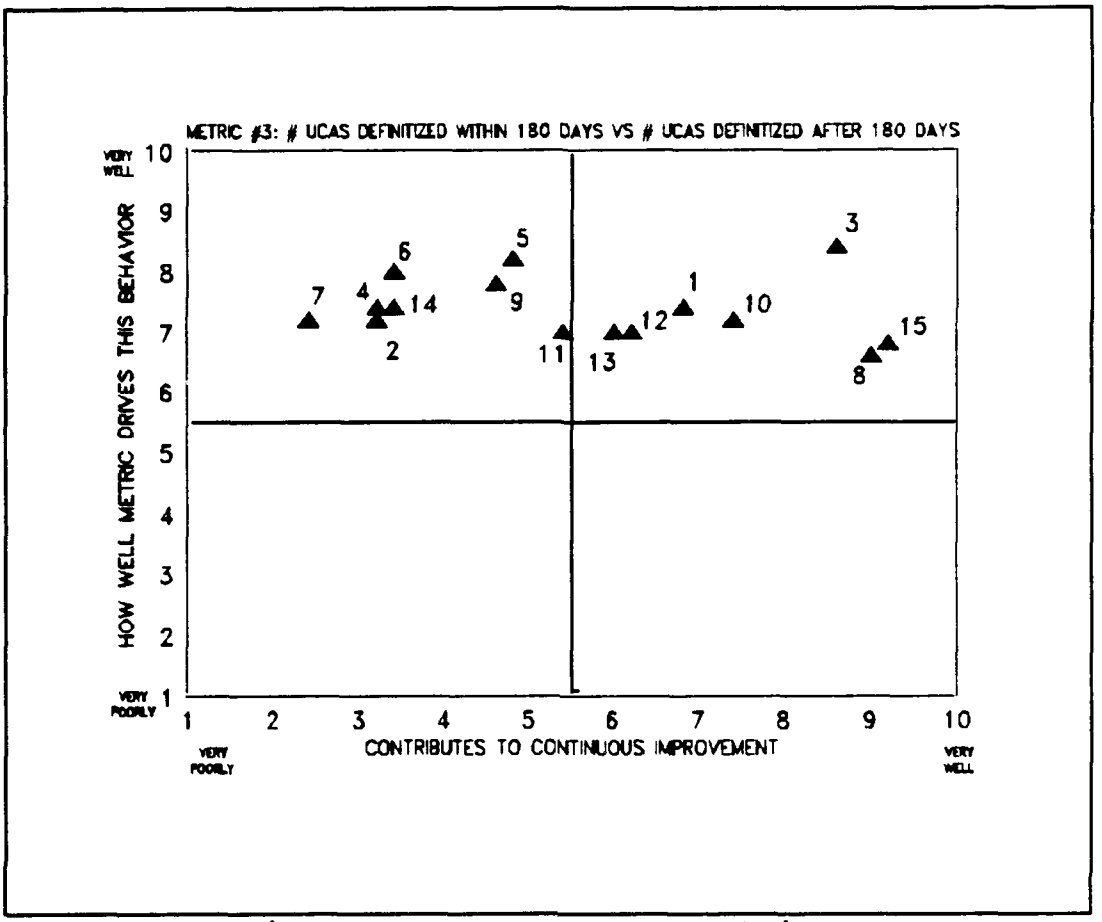


Figure 4-3. Results of Metric 3

the customer. One example of the difference is in terms of "increased reviews." To me, this does

not constitute continuous improvement. It is a form of inspecting in quality, and does not promote "trust" and "empowerment".

This metric fails to focus on eliminating UCAs. It further fails to focus on decreasing the 180 day limit.

Decreasing the use of UCAs is not necessarily good because UCAs are management tools.

There should be complementary metrics in order to focus the behavior better. Decreasing the schedule may be good, but reducing the quality is not. A schedule metric must be combined with a performance (quality) metric.

Analysis. The evaluation group presented results for this metric that were similar to the first two. Figure 4-3 reveals a roughly even split among behaviors that contribute to continuous improvement and those that don't. Again, it appears that this metric is concerned with meeting a requirement while leaving the quality of the work as a secondary consideration. In their rush to meet the 180 day deadline, the SPOs may be driven to conclude agreements that are not actually in the best interests of the government. For example, behaviors two, four, and six actually result in a de-emphasis on UCA quality in order to meet the 180 day deadline. Clearly, the government is negotiating at an extreme disadvantage and may feel forced to conclude any deal to avoid the 180 day deadline. Behavior number seven, in which the contractor actually slows down the negotiating process to force the government to become desperate, supports this assertion. Additionally, this desperation may lead the SPO to shift resources toward resolving the UCA at the expense of other SPO activities. Behavior fourteen is an example of suboptimization in that the behavior being driven meets the goal of definitizing within 180 days but at the expense of other SPO activities. Thus, the net result of this metric might be a hastily reached agreement which is not in the best interests of the government.

On the other hand, this metric is driving some behaviors that do contribute to continuous improvement. This appears to be especially true concerning those that

focus attention on improving the process of negotiating UCAs or, better yet, those behaviors that lead to less UCAs being required to begin with.

Another interesting observation about this metric concerns the spread of individual ratings about whether or not the behaviors are contributing to continuous improvement. It appears the group members had very different views concerning whether increased activity as the 180 day limit approached was positive or negative. A possible reason for this was the absence of any objectives to go along with this metric. Without a common reference from which to start, the group members were relying on individual experience and personal bias. Thus, it seems critical to have metrics linked to objectives if a truly accurate judgment is desired.

Metric 4

The fourth metric evaluated was "Number of project baseline breaches per month". Table 4-5 contains the list of likely behaviors this metric may drive.

TABLE 4-5

LIKELY BEHAVIORS RESULTING FROM METRIC 4
NUMBER OF PROJECT BASELINE BREACHES

1. SPO will work to decrease number of breaches
2. Project managers will interpret their own data as non-breaches

TABLE 4-5 (CONT)

LIKELY BEHAVIORS RESULTING FROM METRIC 4
NUMBER OF PROJECT BASELINE BREACHES

3. SPO project team will set more conservative baselines
4. Baselines written only on those actions with high success possibilities
5. Project team will write fewer baselines
6. Manpower will be shifted to work breached projects
7. Increased project status review/reporting
8. Upper management will help breached project managers
9. SPO will work individual breaches before entire program is breached
10. PMs emphasize the term "informal" in order to explain away breaches
11. Monthly reporting will shift management emphasis to short-term problems
12. Finger pointing increases
13. Problems will be worked according to the probability of causing a breach
14. Easy-to-work breaches will be solved before tough ones
15. PMs will divide work according to number of potential breaches
16. Program Manager focuses on number of breaches rather than the reasons for or severity of the breach

Rating Summary for Metric 4. The combined rating for the fourth metric is illustrated in the composite graph in Figure 4-4. The individual behavior graphs are located in Appendix A.

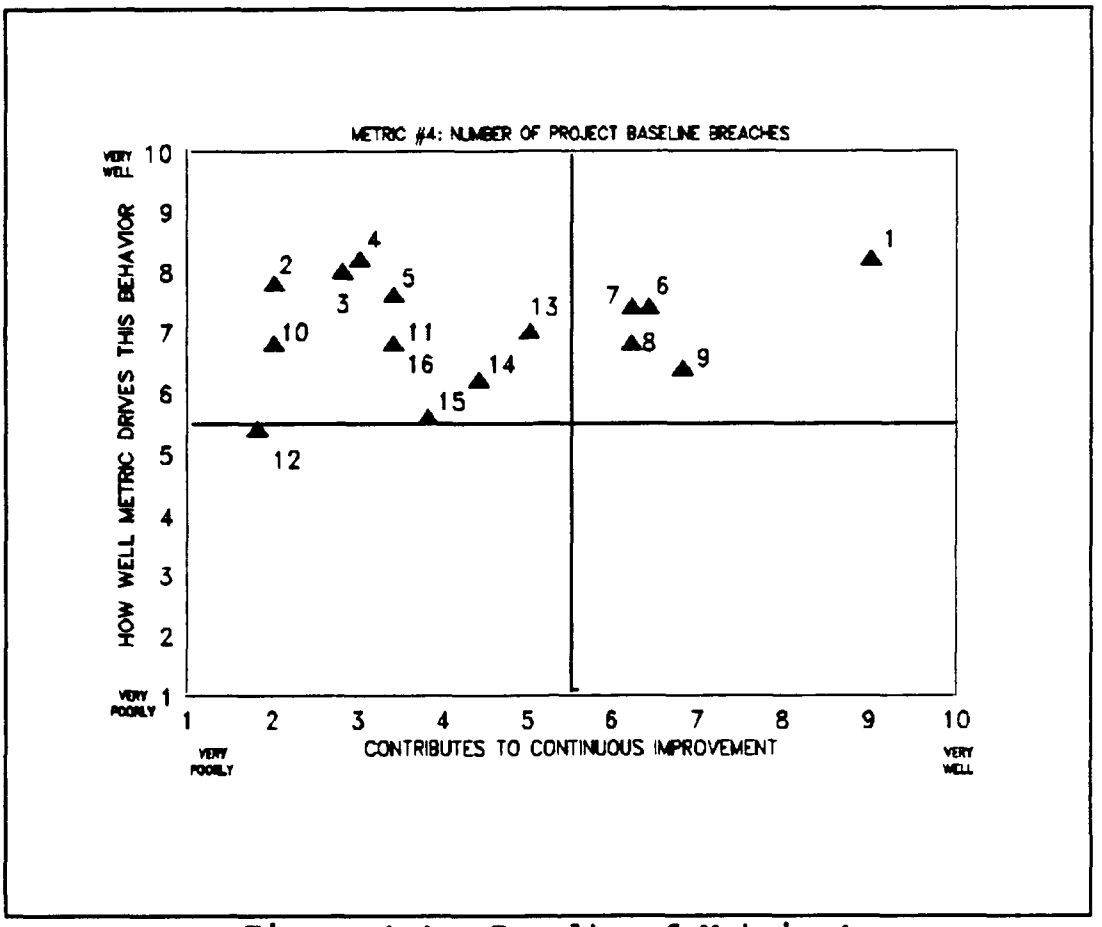


Figure 4-4. Results of Metric 4

Evaluation Group Comments. The evaluation group recorded the following comments about the fourth metric following the rating of the behaviors.

> A decrease in the number of baselines set is not necessarily good. Baselines are important in order to assess program performance. Good baselines are better. Fear of failure is not a good reason not to establish baselines.
>
> My current program is in the middle of requirements definition and baseline formulation. Because of the fear of future breaches, several of the acquisition participants strongly opposed the user creating challenging requirements. Furthermore, they tried several times to get him

to reduce the number of requirements and/or water them down to vague statements.

Analysis. Unlike the three previous metrics, the majority of the behaviors likely to result from this metric do not contribute to continuous improvement. One of the main results of this metric is an increase in fear. For example, behavior sixteen, focussing on the number of breaches rather than the reasons for or severity of the breaches, is a direct result of a fear of having to report a breach. This fear can also lead program managers to attempt to explain away breaches as "inconsequential" since they are not formal breaches. Even worse, this metric may lead to weak baselines being established to ensure they can be met. Behaviors three, four, and five address this issue. The behaviors being driven tend to result in a "watching out for number one" approach to management. Setting baselines that are easily met may support an objective of reducing breaches but, at the same time, may inhibit the accomplishment of the ultimate program goals. Project timelines may be stretched out resulting in delays in fielding the system in question. Excessive spending baselines may be established to hide cost overruns that would have shown up if more realistic baselines had been established. Even worse, if the user is pressured to accept reduced performance baselines, the fielded system may not meet the actual requirements of the operational environment.

Without realistic baselines it would be difficult to identify potential problem areas of the program. Additionally, opportunities for improving cost, schedule, and performance may be missed because of overly conservative baselines. People may be content with simply meeting the goals while ignoring the potential to improve on the easy-to-achieve baselines. Establishing more difficult, yet realistic baselines could provide improved program performance by allowing genuine problem areas to be identified and addressed. A metric that addresses only the number of breaches hinders this process. A better metric might include the causes of the breach. In this way, problems outside the control of the person or persons being measured could be identified. This would remove some of the fear of reporting breaches. While this might increase finger-pointing to find the source of the breach, in the long run it would result in efforts to solve the problem as opposed to hiding the problem with conservative or non-existent baselines. Measuring somebody based on the results of a process outside that person's control is anathema to the concept of total quality management. The fear must be driven out before improvement can take place. As Deming states "No one can put in his best performance unless he feels secure" (11:59).

Finally, it appears that many of the likely behaviors that result from this metric are not actions that will improve the process. In other words, they will not decrease

the number of breaches. An example is behavior six, shifting manpower to work breached projects. Although this is an attempt to fix the breach, it does nothing to prevent future breaches from occurring. Behavior eight is similar. Receiving assistance from upper management after the breach has occurred contributes little to preventing future breaches. The emphasis should be on developing realistic baselines and then managing proactively with the goal of meeting those baselines. This metric emphasizes avoiding tough baselines and working the problems after they occur. This is not consistent with continuous improvement or providing customer satisfaction.

Metric 5

The fifth metric evaluated was "Average variance in days between the scheduled release of Requests for Proposals (RFPs) and the actual RFP release by month". Table 4-6 contains the list of likely behaviors this metric may drive.

TABLE 4-6

LIKELY BEHAVIORS RESULTING FROM METRIC 5
AVERAGE VARIANCE IN DAYS BETWEEN SCHEDULED RELEASE
OF RFP AND ACTUAL RELEASE

1. The SPO will make conservative estimates of scheduled release dates
2. The SPO will concentrate on meeting the schedule rather than preparing a good RFP
3. Determine the components of the RFP cycle

TABLE 4-6 (CONT)

LIKELY BEHAVIORS RESULTING FROM METRIC 5
AVERAGE VARIANCE IN DAYS BETWEEN SCHEDULED RELEASE
OF RFP AND ACTUAL RELEASE

4. SPO increases reviews of proposal development status
5. SPO will develop a good process for RFP development
6. There will be increased visibility on schedule intensive proposal development tasks
7. The SPO will pressure coordinating offices to shorten review
8. Manhours will be moved from other functions of the program to work on the RFP/ECP
9. The SPO will tend to used "canned" RFP packages
10. SPO/Contractor team will perform joint RFP development
11. SPO will try to reduce the number of RFPs
12. RFP prep priorities will be based upon ability to meet schedules
13. The SPO will reduce the number of offices required to coordinate on the package
14. The RFP team will be comprised of functionals to write the package vs only one functional doing the writing
15. The SPO will try to accelerate all RFP preps as much as possible
16. SPO will not wait on a clear requirements definition from the user in order to meet the RFP schedule
17. The SPO will schedule CCBs on an as-needed basis versus a routine basis
18. SPO will try to get an informal jump on RFP prep before official approval is received
19. "Shotgun" proposals to each office vs. in-turn routine correspondence coordination
20. SPO will do a large number of small ECPs

TABLE 4-6 (CONT)

LIKELY BEHAVIORS RESULTING FROM METRIC 5
AVERAGE VARIANCE IN DAYS BETWEEN SCHEDULED RELEASE
OF RFP AND ACTUAL RELEASE

21. SPO will be less likely to take on multiple RFPs simultaneously

Rating Summary for Metric 5. The combined rating for the fifth metric is illustrated in the composite graph in Figure 4-5. The individual behavior graphs are located in Appendix A.

Evaluation Group Comments. The evaluation group recorded the following comments about the fifth metric following the rating of the behaviors.

> There was some confusion over the meaning of "SPO reviews." Some thought it meant SPO Director reviews and some thought it meant lower level reviews with the SPO.
>
> The metric should be specific when describing RFP vs ECP, etc. Additionally, there is significant confusion over the levels of reviews associated with the metric. It appears to be very important to specifically identify who uses the metric and how the metric is to be employed along with the objective.
>
> This metric needs to be changed to include ECPs as well as RFPs. In most cases, the program office will do multiple ECP cases and only one RFP. Unless you find yourself in a basket SPO, then the SPO may do multiple RFPs during its lifetime. But each individual project within the basket SPO will perform only one.
>
> Interesting enough, the behaviors with regards to meetings and reviews should be broken to include meetings/reviews at the higher levels, and those at the lower levels. It would be interesting to

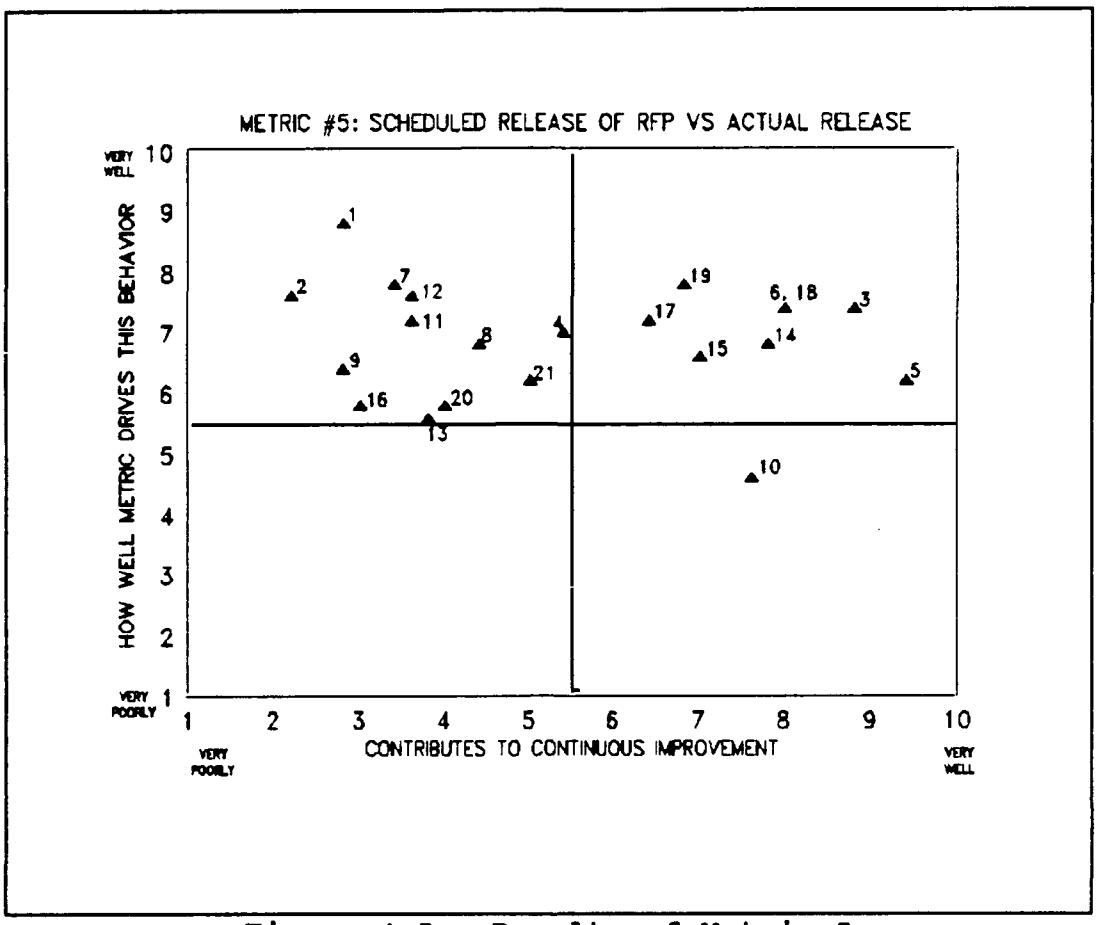


Figure 4-5. Results of Metric 5

note the differences in how the requirement for these types of reviews differ with respect to rank or job title.

This metric takes the focus off reducing the number of days required. The metric should be viewed in conjunction with another metric that looks at both the number of days to process and the variance from planned. An important factor is the target number of days and how satisfactory is it to the customer.

It may be difficult in the first place to accurately forecast the time it takes to develop the RFP. We can only go by past experiences. There are plenty of outside factors such as requirements creep that extend the RFP process. The long-term goal should be to develop a generic process (with a timeline) that will give you the

highest quality product. Does the metric do that?

Analysis. The cloud chart indicates a general consensus among the group regarding the likelihood of the behaviors associated with this metric. All but one of the triangles are located in the upper two quadrants. As was previously mentioned, this indicates these behaviors are likely to occur in response to this metric. Specifically, triangles in quadrant two reflect behaviors likely to be driven by the metric but which do not lead to continuous improvement. Behaviors one, two, seven, eleven, and twelve were considered to have the most significant negative impact to continuous improvement. These five behaviors are located in the upper left portion of quadrant two.

These behaviors appeared to be driven by the pressure to release an RFP and meet the projected schedule. Creating conservative schedules, concentration on schedule vs. the RFP, shorter reviews, focusing on RFPs closest to the scheduled release date (while ignoring others), and reducing the number of RFPs are all negative behaviors associated with attempts to meet a schedule.

It is difficult to determine a realistic schedule for RFP development. Many outside factors such as changing requirements (by both user and the SPO), inter-SPO project priorities and team composition can affect a schedule in a variety of ways. Developing a specific/generic RFP

development process within the SPO could lead to the creation of more realistic schedules and take some of the pressure off the program managers.

The cloud chart scatter concerning the behaviors affect on continuous improvement indicate that this metric is one that could use some improvement. The improvement could come through modification of the metric or by combining it with other metrics to make it more effective and eliminate some of the negative behaviors. Twelve of the behaviors were considered to be detrimental to continuous improvement and nine of the behaviors were considered to be positive to improvement.

There was very little variance among the group concerning each behavior's ratings. In other words, there were few instances where two or three of the group members rated a behavior an "8" and two or three of the others rated the behavior a "2." The only polarization occurred in behaviors four and seven. These behaviors focused on reviews. There was disagreement "evenly split" among the group concerning the value of many versus few reviews and long versus short reviews.

Metric 6

The sixth metric evaluated was "Number of contractor data submittals due versus those actually received on a monthly basis". Table 4-7 contains the list of likely behaviors this metric may drive.

TABLE 4-7

LIKELY BEHAVIORS RESULTING FROM METRIC 6
NUMBER OF CONTRACTOR DATA SUBMITTALS DUE VS. RECEIVED

1. SPO engineering data management officer (EDMO) will be more concerned with data submittal dates than with data content
2. SPO focus on contractor versus SPO review/approval
3. Contractor will be more concerned with data submittal dates than with data content
4. Weekly status reports will be generated and tracked by the SPO data manager
5. Contractor will request extensions of due dates until later in the program
6. SPO personnel will not review data; they'll just track its submittal
7. Contractor will request that SPO reduce data requirements
8. SPO will conduct periodic reviews of essential data requirements for additions/deletions
9. SPO will perform in-depth tracking of the contractor data management system
10. Contractor will shift manpower to work on late data submittals
11. Contractor will find excuses for late data deliveries
12. SPO will seek to do up front work with the contractor to improve submittals
13. Status of data submissions will be briefed at program reviews
14. Contractor will develop a process to get quality data out in a timely manner

Rating Summary for Metric 6. The combined rating for the sixth metric is illustrated in the composite graph in

Figure 4-6. The individual behavior graphs are located in Appendix A.

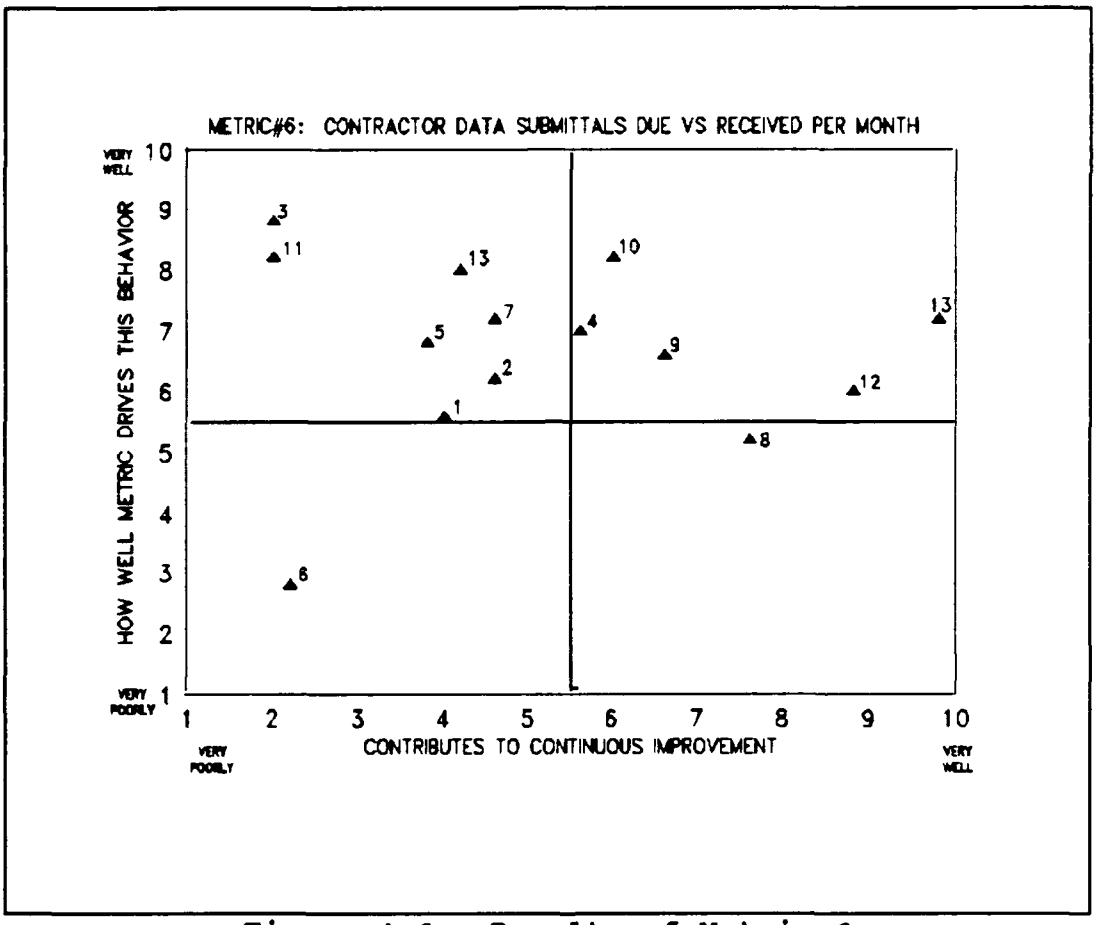


Figure 4-6. Results of Metric 6

Evaluation Group Comments. The evaluation group recorded the following comments about the sixth metric following the rating of the behaviors.

> This metric falls short of measuring the entire configuration/data management topic. It does not measure the requirements for a quality product, nor the Government's ability to measure that quality in a review cycle.

> It is difficult to set up a metric that will measure contractor performance unless the government has some sort of contractual control over the contract for that metric -- like an award fee. If not, you get into a situation where the contractor couldn't care less what your metric is.
>
> This metric also needs to be viewed in conjunction with other metrics to measure quality. This metric in isolation could result in sub-optimization or sub-allocation of resources in the SPO or with the contractor.

Analysis. The cloud chart indicates all but two of the behaviors identified by the group are likely to be driven from the use of the metric. And once again, many of the behaviors were rated as having a negative affect on continuous improvement. Behaviors three and eleven were rated as being the most likely behaviors and having the most negative impact on continuous improvement. Both behaviors focus on the contractors. Behavior three has to do with the contractor being more concerned with data submittal dates than with data content. Behavior eleven focuses on the contractor finding excuses for late deliveries.

The fact that these two "bad" behaviors are the most likely to occur **and** lend themselves least toward continuous improvement, indicates that this metric needs improvement. In addition, the wide spread of good and bad behaviors identified with this metric point to further troubles with its use. However, as one of the group members suggested in his comments, this metric could prove effective if coupled with another metric with a focus on quality.

In all of the group comments, there was a common concern about the affect the metric would have on contractor data quality. When there is no link with the quality of a product whose schedule is being tracked, inferior products can result. In other words, the contractor could be so concerned with meeting schedule that he'll cut back on quality. The result is an on-time product that is technically unacceptable. As mentioned previously, if the schedule metric was used in tandem with a quality measurement, it could prove to be more effective.

Metrics one and six were similar in one aspect. Both metrics focused on a product delivery **required** versus the number actually delivered. It appears that those behaviors that focus on meeting a specific target (i.e. goal or quota) do not lead to continuous improvement. Conversely, those behaviors that are focussed on improving the process used to meet the goal do lead to continuous improvement. There were more of these type behaviors driven by metric one than by metric six. A key reason may be that metric six is measuring outside the control of the SPO. Therefore, most of the behaviors being driven will do little to improve an external process. One of the comments made by an evaluator addressed this concern. Essentially, the evaluator was concerned that this metric is measuring something that is essentially outside the control of the SPO. Unless the SPO has some contractual control over the schedule of data submittals, this may well be a useless metric. The

contractor can submit the data whenever it is convenient and do so with impunity. This will lead to frustration if someone within the SPO is held accountable for the contractor's performance. In this case, the SPO definitely requires a strong contractual incentive to ensure the contractor is meeting the needs of the government.

Metric 7

The last metric evaluated was "Number of Government Furnished Equipment deliveries provided vs. the number actually required on a monthly basis". Table 4-8 contains the list of likely behaviors this metric may drive.

TABLE 4-8

LIKELY BEHAVIORS RESULTING FROM METRIC 7
NUMBER OF GFE DELIVERIES PROVIDED VS. NUMBER REQUIRED

1. SPO will increase review of GFE delivery status
2. SPO continually scrubs contractor requirements
3. Government will work to match actual with required
4. GFE deliveries will be prioritized by availability instead of importance
5. Item managers will pressure the SPOs to reduce GFE requirements
6. Item managers will attempt to increase GFE stocks
7. SPO will increase acceptance of poor quality GFE in order to avoid delivery delays
8. Government will work GFE deals off-line until it's sure it can deliver

TABLE 4-8 (CONT)

LIKELY BEHAVIORS RESULTING FROM METRIC 7
NUMBER OF GFE DELIVERIES PROVIDED VS. NUMBER REQUIRED

9. SPO will increase attention to GFE to ensure GFE producers deliver products on-time
10. The SPO will develop an efficient GFE production and delivery process
11. Contractors will increase incoming GFE inspections
12. SPO will tend to reduce the amount of GFE on a project
13. SPO will use quasi-black market tactics to get GFE

Rating Summary for Metric 7. The combined rating for the seventh metric is given in the composite graph in Figure 4-7. The individual behavior graphs are located in Appendix A.

Evaluation Group Comments. None.

Analysis. The group was initially confused as to whether the metric was focusing on common stock type GFE or Government project type GFE. For example, previously purchased special test equipment being used on an acquisition program is different than an element (a black box for example) in an acquisition project that is being managed and purchased by the Government. The team decided that both common stock GFE and project-type GFE would be acceptable to rate. Incidentally, this was the only metric which rates how the Government performs, **not** the contractor.

Curiously, most of the behaviors this metric drives appear to lead to continuous improvement. Perhaps this is a

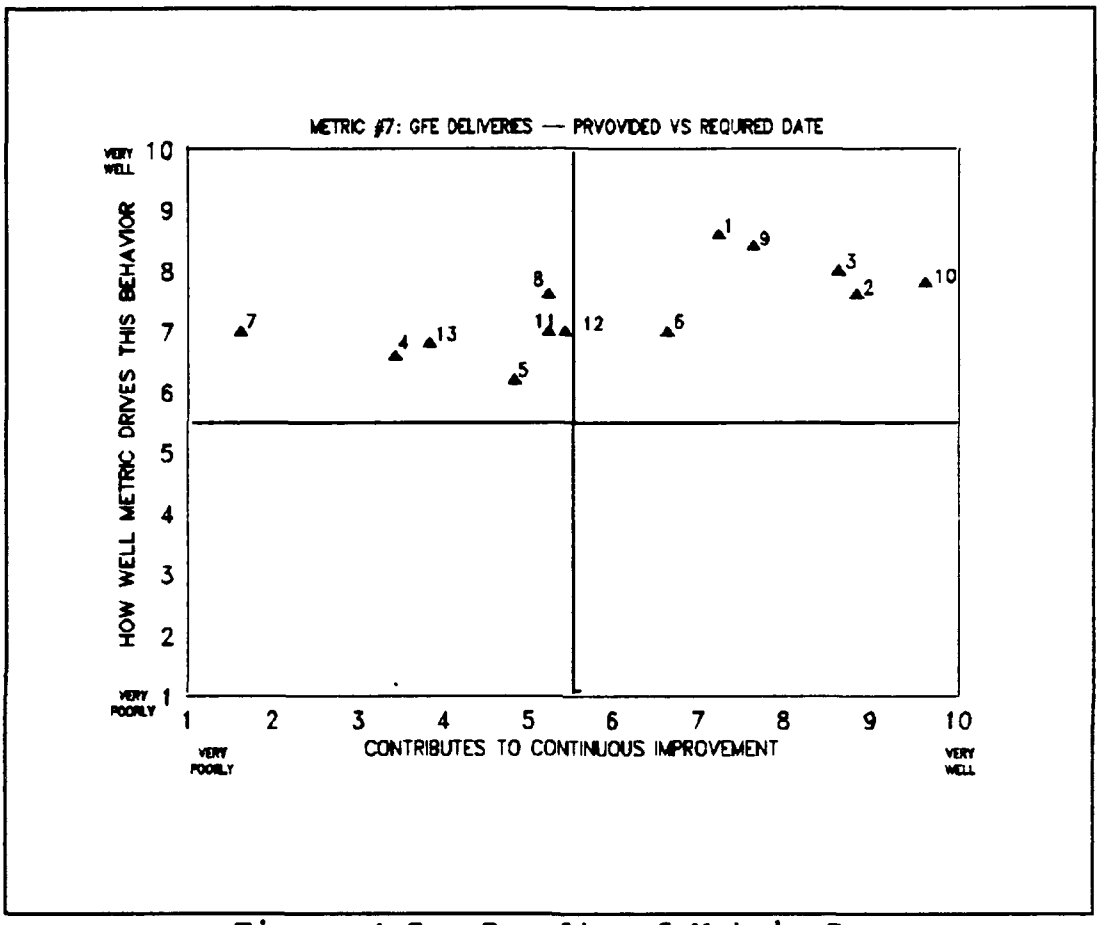


Figure 4-7. Results of Metric 7

tendency to see the worst in the behavior of others (i.e. contractors) while not holding the government to the same high standards. Analysis of additional metrics that measured government performance would prove interesting.

The cloud chart shows all behaviors identified were likely to be driven by the metric. In addition, the scatter showed a tendency toward quadrant one of the chart. This

indicates the metric is well on its way to driving behaviors that lead to continuous improvement. Although only six of the thirteen behaviors were thought to lead to continuous improvement, four of the remaining seven were close to the "middle ground" in terms of continuous improvement. Again, these behaviors are those that focussed on the process rather than the goal of the process. Behavior ten is an excellent example of this behavior. The only "outlier" was behavior number seven. This behavior focused on the SPO accepting poor quality GFE in order to avoid delivery delays. While it had the lowest rating in terms of continuous improvement, it was also among the lowest in likelihood of being driven by the metric. The fact that this was the last metric evaluated could explain why there were no evaluation group comments (fatigue factor).

V. Conclusions And Recommendations

Conclusions

As a result of this thesis effort, the research team formed eight major conclusions.

1. A single metric may need to be integrated with others to be truly effective.

2. Metrics can lead to sub-optimization in the functional areas within a SPO.

3. Behaviors that focus on exploring and improving processes promote continuous improvement. Behaviors that focus on goals, quotas, and the end result usually do not lead to continuous improvement.

4. The field of metrics is a challenging area of study because of the unique features of not-for-profit organizations.

5. In order to be fully understood and correctly used, metrics need to be coupled with an objective.

6. If the metric focuses on an activity the SPO has no control over, it shouldn't be used.

7. Too many metrics can be detrimental to the program office.

8. SPOs should consider using Group Support Systems (GSS) to develop their own internal metrics.

A single metric might not, by itself, lead to continuous process improvement. In other words, it might not be "fair" to rate some metrics' effect on behavior and continuous improvement in a vacuum. A combination of integrated metrics could help to further continuous process improvement. An example of this would be metric six.

Metric six dealt with contract data submissions due vs. those actually received. This focus on schedule is important to measurement. However, too much focus on meeting schedule could result in neglecting the quality aspects of the data products. When a SPO receives an inferior data deliverable from a contractor (a software system specification for example), it could take more time modifying and coordinating the substandard document with the contractor (and within the SPO) than it would to extend the contractor's due date. By coupling the schedule metric with a metric focusing on the quality aspect, many of the negative behaviors associated with this metric could be eliminated. The challenge is for the SPOs to identify the metrics that would be effective toward improving the process. Metrics are receiving increased emphasis in program management. And although metrics are critical to improving quality in organizations, it's important to keep in mind that misuse of metrics can have a negative effect on organizations.

Metrics may lead to sub-optimization of the functional areas versus the greater well being of the SPO or the government. For example, is it proper for a metric to influence the behavior to get a contract modification on contract as quickly as possible? Does this contribute to continuous process improvement? It depends on the situation and the motivation of the people involved. The contracts organization can point to its accomplishment as a matter of pride (and brief it as successfully meeting the metric's goal). If the goal is to get a contract awarded as quickly as possible, the metric does lead to

continuous improvement. However, if the goal is get on contract as quickly as possible while at the same time negotiating a "reasonable" not-to-exceed price, the metric may not influence behaviors that lead to continuous improvement. The Government may pay the price for this behavior in higher program costs as a result of the haste to put the modification on contract. As you can see, a single metric (especially in the Government) can motivate a number of people to behave in different ways. This points out one of the challenging aspects of metrics.

Each of the metrics that were evaluated produced some behaviors that contributed to continuous improvement and some that did not. Almost without exception, the behaviors that did not promote continuous improvement were those that focussed on numerical goals or quotas, or otherwise concentrated on meeting the end requirement of the process being measured. These behaviors usually ignored the process itself. This is a big mistake from an improvement standpoint. Only after examining the process and eliminating the special causes of variation can improvement be achieved. Even more important, attempting to circumvent the normal process to meet numerical goals may actually worsen the outcome of the process. An example of this is the failure to establish tough, realistic baselines in favor of easily-met baselines or no baselines. The goal of no baseline breaches is easily reached in this case, albeit at the expense of the program as a whole encountering real schedule slippage, cost overruns, and performance deficiencies.

On the other hand, process-focussed behaviors generally lead to continuous improvement. This is because real improvement can only be obtained from understanding and improving a stable process. A good example of this a behavior driven by metric three, "Number of UCAs definitized within 180 days vs. the number definitized after 180 days". The fifteenth behavior results in the contractor and the SPO working more closely during contract negotiations to prevent the need for UCAs. Instead of attempting to deal with UCAs within the time limit of 180 days, this behavior is directed toward eliminating them by perfecting the beginning of the contracting process. Responding to the challenge of developing metrics that drive process-oriented behavior will be a demanding, but ultimately valuable endeavor.

For the acquisition program offices, metric development is made even more challenging due to the fact the focus isn't on the same "bottom line" -- profit motivation -- as in the private sector. Defining the mission of Air Force acquisition in "quality" terms is difficult indeed. The contractor probably knows the budget for the program. The program offices make no profit. Any savings made from contract negotiations will probably be used for another project (or program). What is the product? It's hard to quantify. Because private sector metrics can be quantified more easily, the Government's challenge is made even greater. In addition, metrics need to be directed at activities the Government has control over.

If the Government doesn't have some sort of contractual control over the activity the metric focuses on (like an award

fee), the contractor could have very little incentive to pay attention to the metric. If the metric focuses on an activity the program office can't control, it's useless and shouldn't be used. There is a danger in developing metrics just because their in "style." And there is also a danger of developing too many metrics for an organization.

The team collected over 300 internal metrics among five different program offices. The literature review indicated that too many metrics can be harmful. The organization can lose sight of its mission. So much time is spent collecting, tracking and analyzing the metrics that the "real" work necessary to accomplish the mission of the organization is not being accomplished. The team felt that too many metrics were being used within the SPOs. In addition, most of the metrics lacked objectives. This
lack of objectives makes the quest for continuous improvement even more difficult.

The majority of the metrics collected by the team did not have an objective associated with them. This could have contributed to some of the variance in ratings. The majority of the group members continually asked about the goal/objective the metric was supposed to be linked with. Because the study concerned actual metrics collected (that didn't contain an objective), the group was forced to rate the metrics according to their own unique perceptions. As previously stated, a good metric must be linked to an objective if it is to lead to improvement.

Finally, the team recommends that SPOs use Group Support Systems (GSS) to develop more effective internal SPO metrics. In a process that would essentially take place in the reverse order of this research, an acquisition objective could be selected for evaluation. An evaluation group from the SPO (through a GSS software package) could brainstorm the positive behaviors that would need to be driven in order to meet the objective. The group could then brainstorm a metric, (or better yet a group of metrics), that would drive these behaviors. Finally, they could evaluate the candidate metrics using cloud charts similar to those used in this study.

Metrics can significantly improve the government's managerial processes. There is a great deal of interest and belief that TQM can be applied in an effective manner. The challenge is developing metrics that are appropriate for use in government acquisition offices.

Recommendations For Further Research

The research team has four recommendations for further research that resulted from this project. First, additional types of metrics should be evaluated. Second, the metrics should be prioritized as to the most common among the SPOs. Third, the metrics' presentation methods should be analyzed. And finally, the GSS should be used in a thesis project to develop and/or compare candidate metric development approaches.

Another research team could investigate additional types of metrics. Of the three critical areas of program management,

cost, schedule, and performance, this study only focused on schedule. As was previously stated, the team collected over 300 internal metrics from the five SPOs selected for the study. Cost and performance are metric areas that are also extremely important in the acquisition process. An evaluation of these metrics would be advantageous to ASC and to the Air Force as a whole. Metrics are certain to play a key role in the restructured Air Force. The Air Force acquisition metrics development process is still in its infancy. Continued research in this area can only help the Air Force' progress in the metrics arena and its focus on quality management.

The team encountered a variety of different metrics from the SPOs identified for the research. A study conducted to prioritize the metrics used among the SPOs could prove interesting. For example, it would be interesting to note how many program offices track APDP certifications versus undefinitized contract actions. Since Air Force Acquisition metrics are in their infancy, it might be helpful to analyze their current direction. This research could lead to a generally agreed upon list of core metrics that should be used in all SPOs. Such a list would prove extremely valuable to SPOs which are the still in the initial stages of developing internal metrics.

While the team encountered a variety of metrics among the program offices, it also encountered a number of similar ones. It was interesting to note how metrics measuring the same thing were graphically presented in different manners. For example, a typical product delivery metric was presented by bar charts

(vertically), bar charts (horizontally), XY charts with six month moving averages, XY charts without six month moving averages, textually (i.e no graphics) and in a number of other forms. In addition, the symbols used in the presentation appeared to affect the interpretation of the metric. Each SPO had a unique preference for the graphical presentation used for the metrics. For this reason, the schedule metrics selected for this study were textually "sterilized" for the evaluation group's analysis.

However, it would be interesting to note if any of the metrics graphical presentation styles would be preferred over others. In other words, is one style "less busy" and more easily understood by a manager. If so, maybe it should be used (vice others) in the presentation of this specific metric in management reviews.

Finally, a research team could use the GSS to develop metrics development approaches in much the same way as was mentioned in the previous section of this chapter. The advantages and disadvantages of different developmental approaches could be analyzed. A "polished" metric development approach researched by AFIT thesis students could prove valuable and aid the acquisition program offices in a large way.

Summary

Metrics proved to be a very challenging area of research. Because the interest and focus of the subject is so new, literature on the subject was limited. In addition, metrics, by their very nature can be wide and varied. Often times nebulous

and abstract, they can be interpreted in a number of different ways. Although difficult to develop and agree upon, they can be an extremely effective means of improving quality within organizations. While the difficulty of developing metrics in the governmental sector should be noted, this should not be a reason to abandon the attempt to implement a comprehensive system of efficient, integrated metrics. More work needs to be accomplished in this area. Additional research can only help the process.

Appendix A. Individual Behavior Ratings

Metric 1 Behavior 1. Increased emphasis on delivery schedule requirement/reviews.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	1	2	-	1	7.60
How well does behavior contribute to CI?	-	-	1	-	-	-	1	1	-	2	7.60

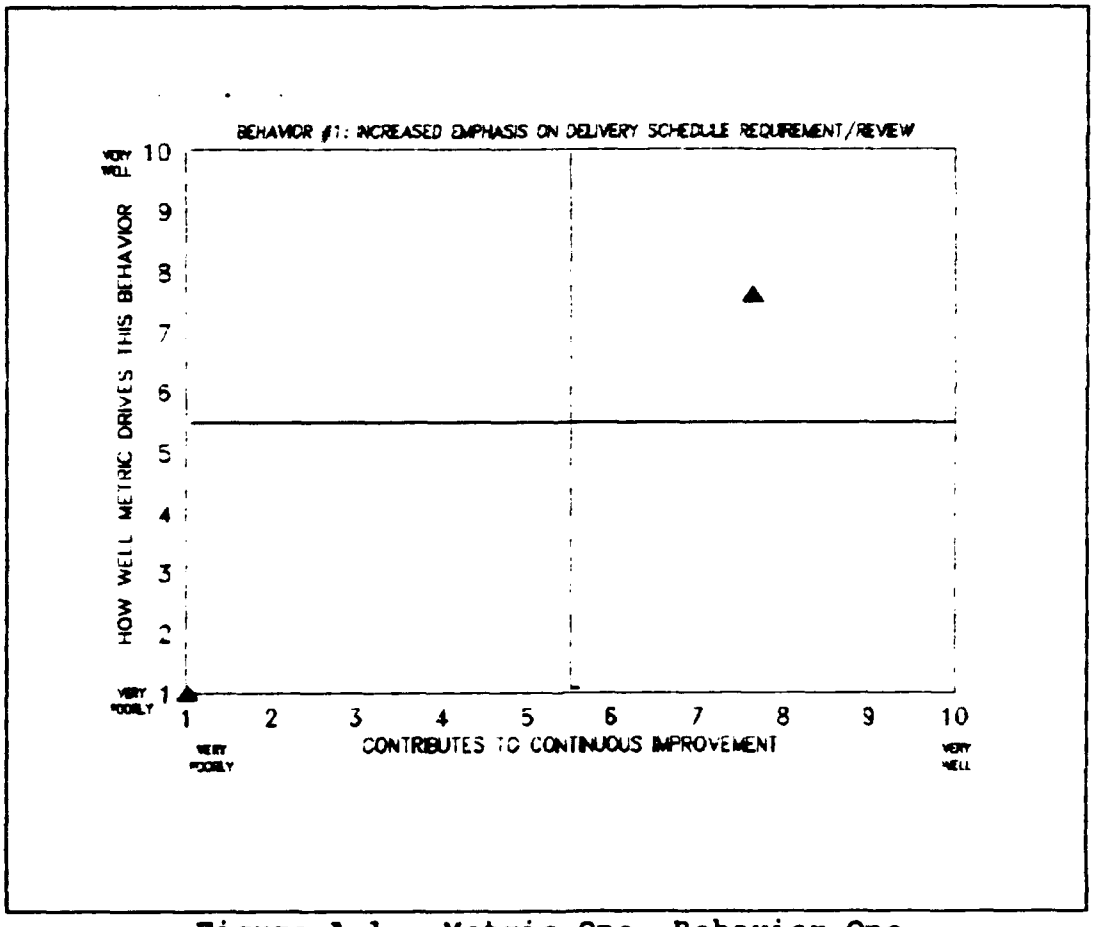


Figure A-1. Metric One, Behavior One

Behavior 2. People will increase deliveries near the end of the month.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	2	1	1	-	7.40
How well does behavior contribute to CI?	-	-	3	1	-	-	-	-	-	1	4.60

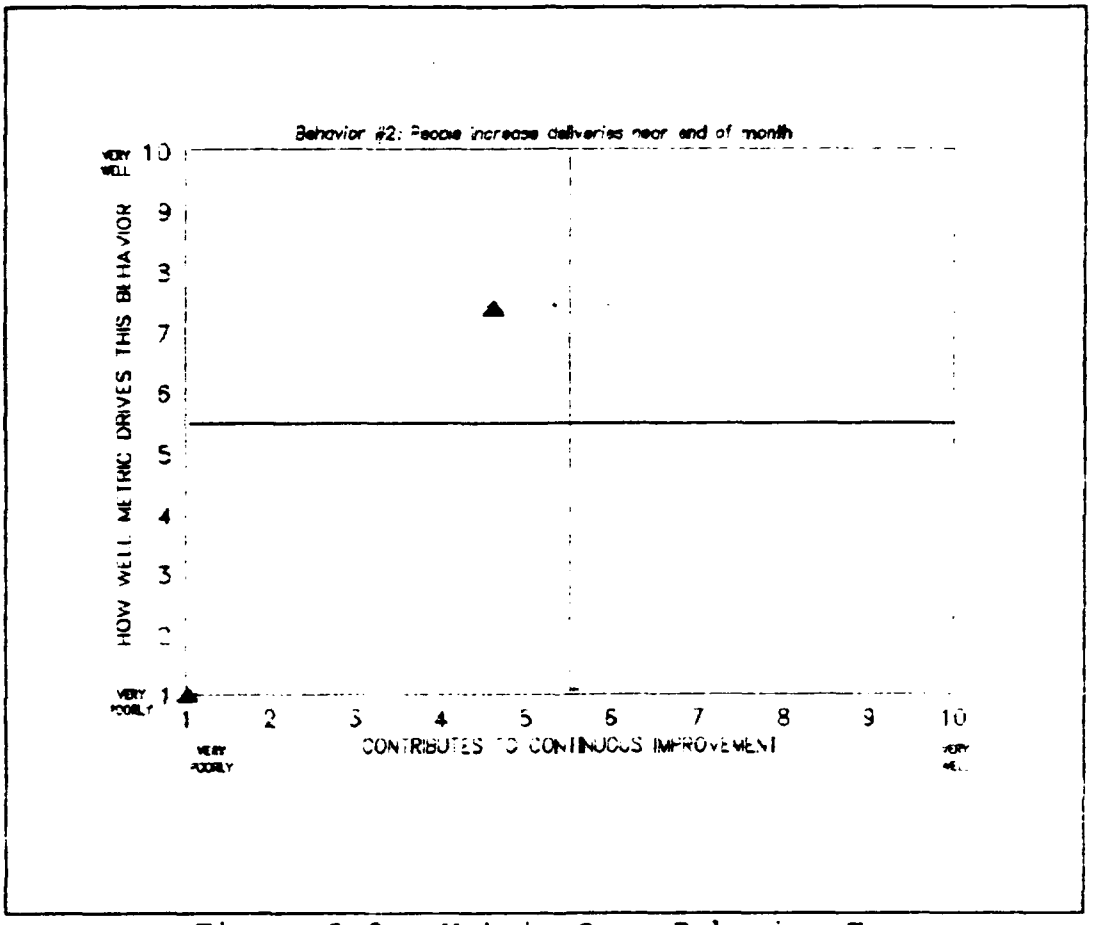


Figure A-2. Metric One, Behavior Two

Behavior 3. Increased emphasis on project planning.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	-	1	-	1	-	2	7.40
How well does behavior contribute to CI?	-	-	-	-	-	-	-	1	1	3	9.40

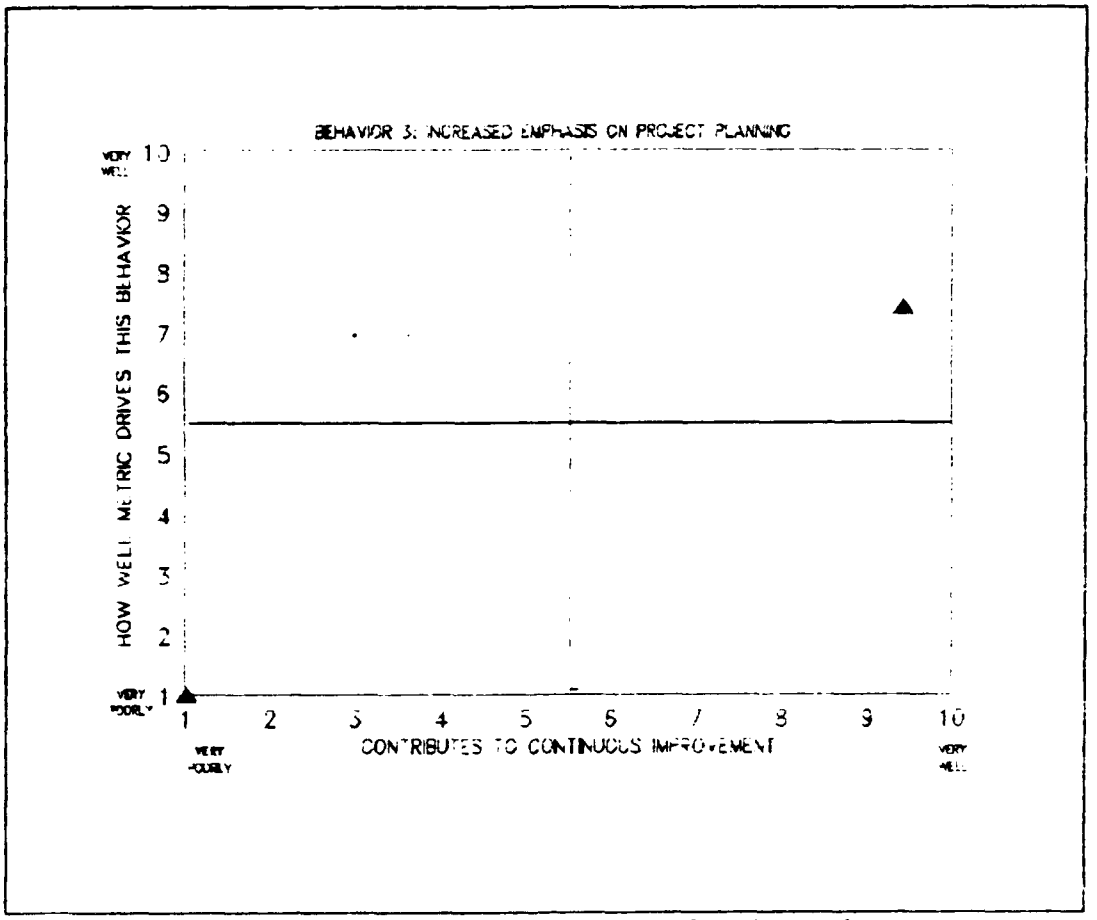


Figure A-3. Metric One, Behavior Three

Behavior 4. People will seek a high quantity requirement if they're good workers.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	2	1	-	1	-	1	-	-	-	3.80
How well does behavior contribute to CI?	1	2	1	-	-	1	-	-	-	-	2.80

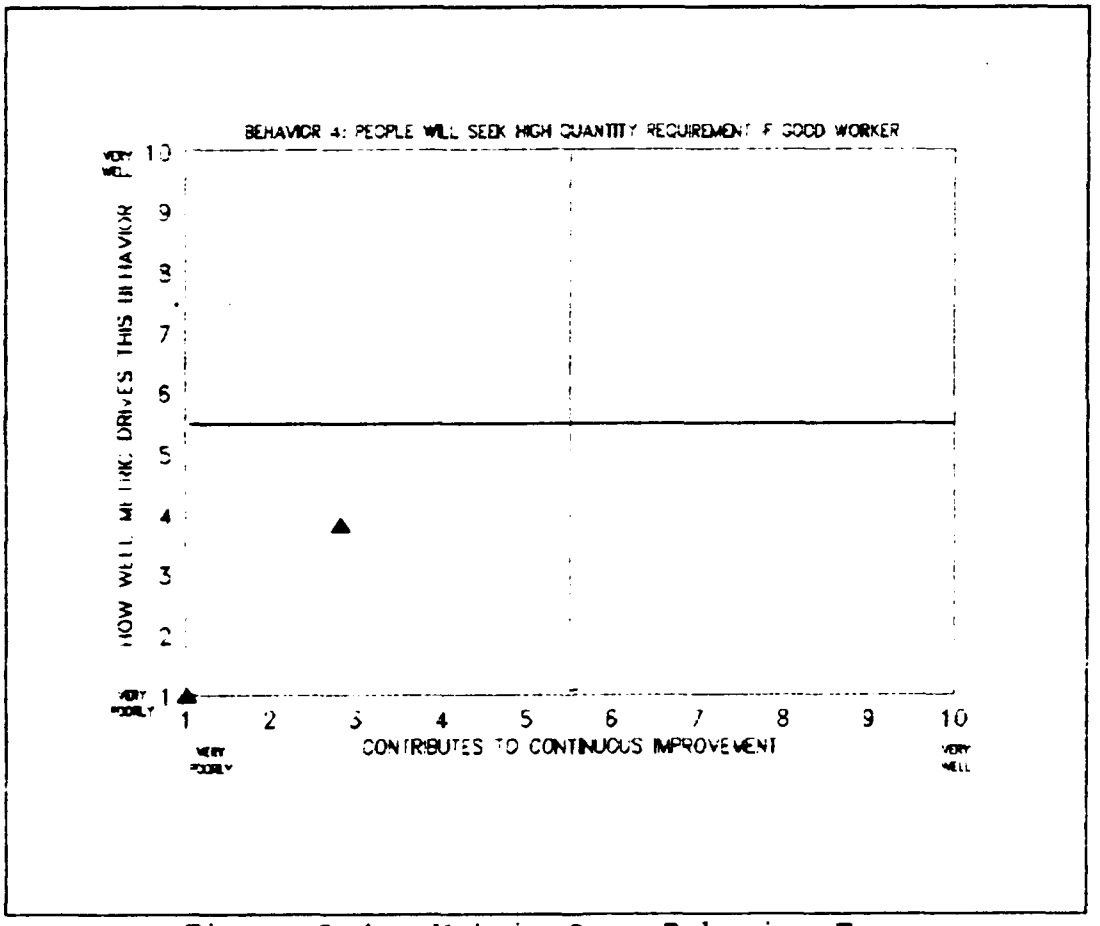


Figure A-4. Metric One, Behavior Four

Behavior 5. People will change the schedule deliveries to meet output

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	1	2	1	-	7.40
How well does behavior contribute to CI?	1	-	1	-	1	-	2	-	-	-	4.60

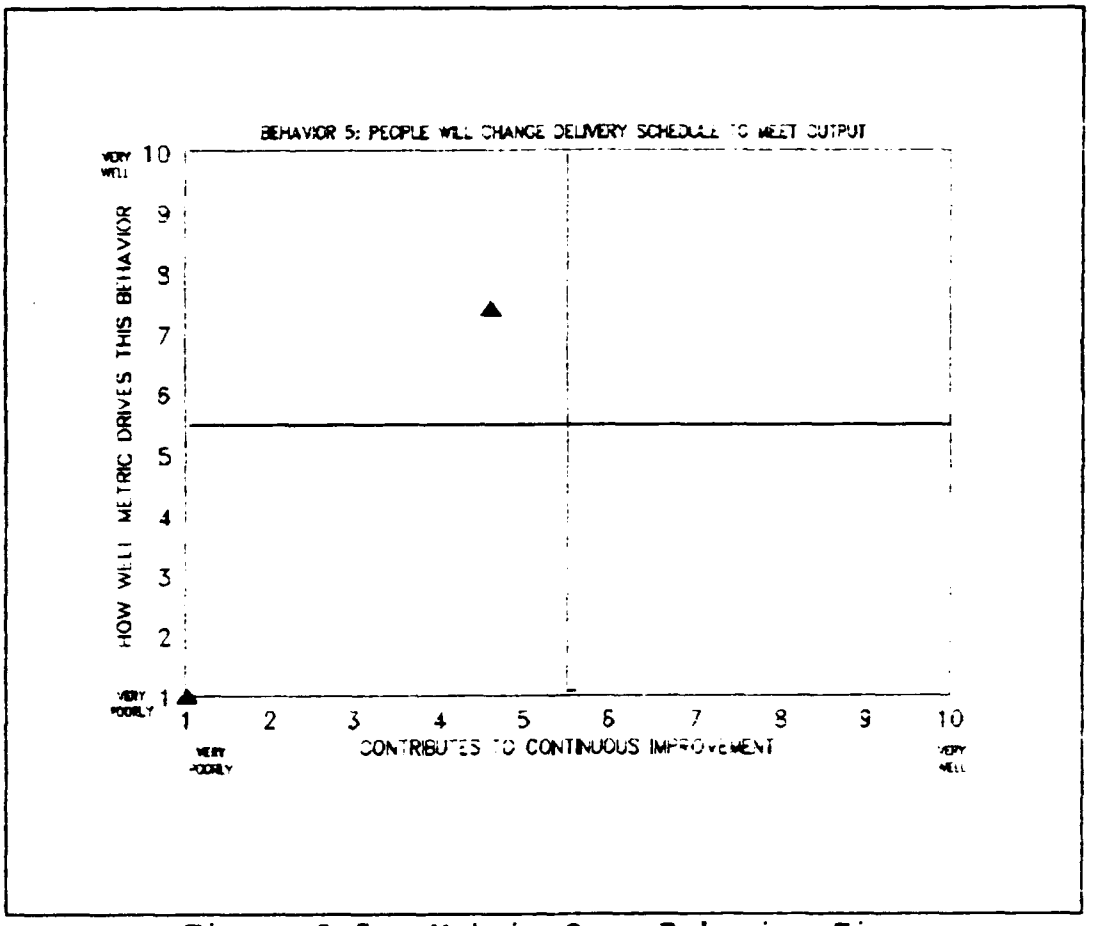


Figure A-5. Metric One, Behavior Five

Behavior 6. Increased emphasis on critical supplier schedules.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	1	3	-	1	8.20
How well does behavior contribute to CI?	-	-	1	-	-	-	2	-	-	2	7.40

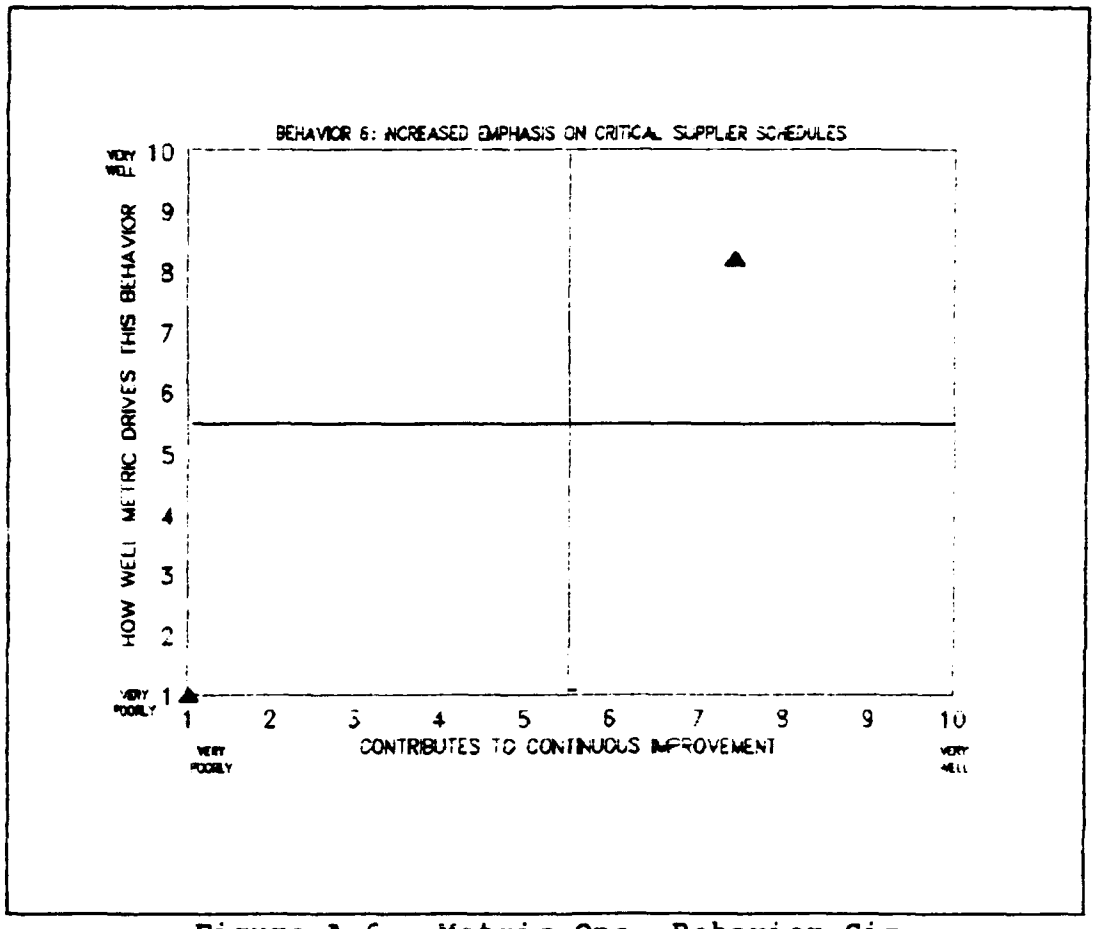


Figure A-6. Metric One, Behavior Six

Behavior 7. People will increase the amount of overtime to meet required deliveries

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	2	-	7.20
How well does behavior contribute to CI?	-	1	3	-	-	1	-	-	-	-	3.40

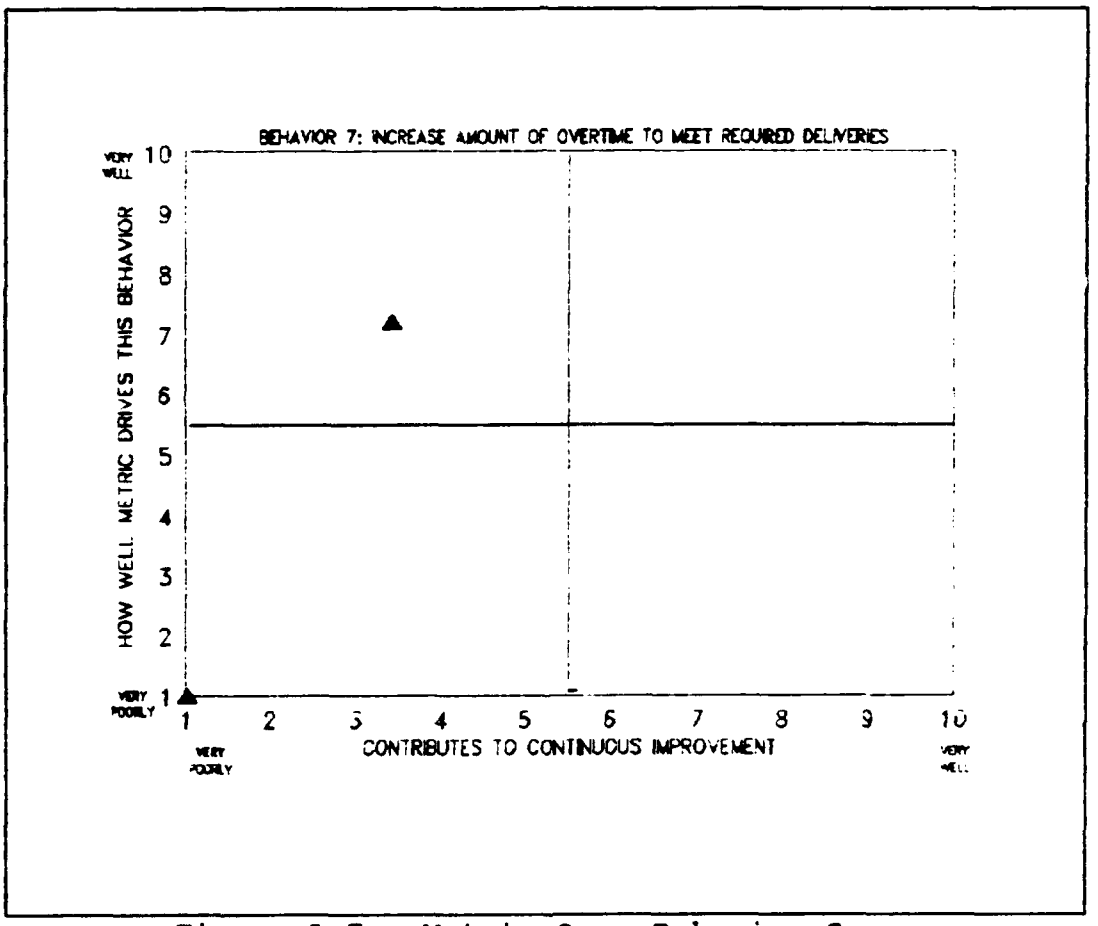


Figure A-7. Metric One, Behavior Seven

Behavior 8. Increased emphasis on critical assembly processes (schedule sensitive).

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	1	2	-	1	7.60
How well does behavior contribute to CI?	-	-	-	-	-	-	-	2	1	2	9.00

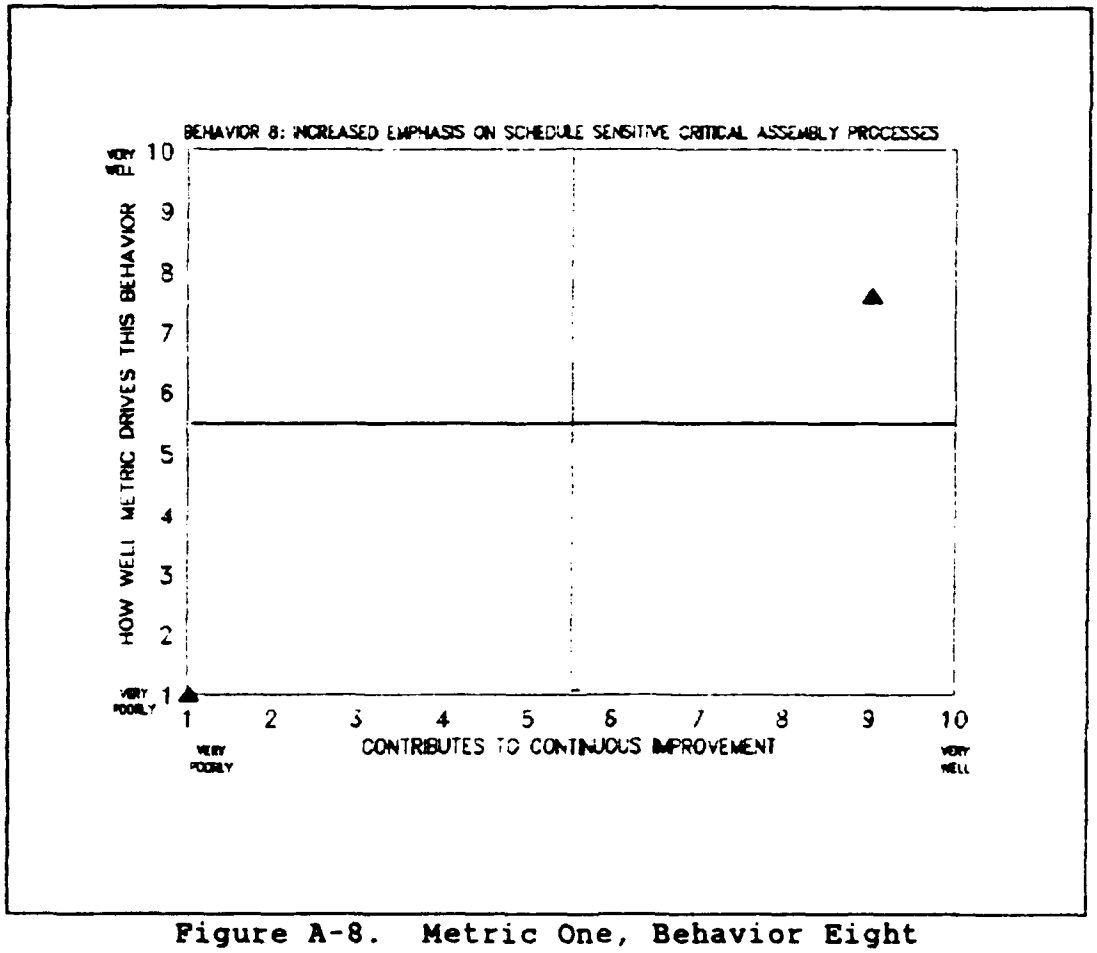


Figure A-8. Metric One, Behavior Eight

Behavior 9. People will increase finger pointing.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	2	-	1	-	-	1	1	-	5.60
How well does behavior contribute to CI?	3	1	1	-	-	-	-	-	-	-	1.60

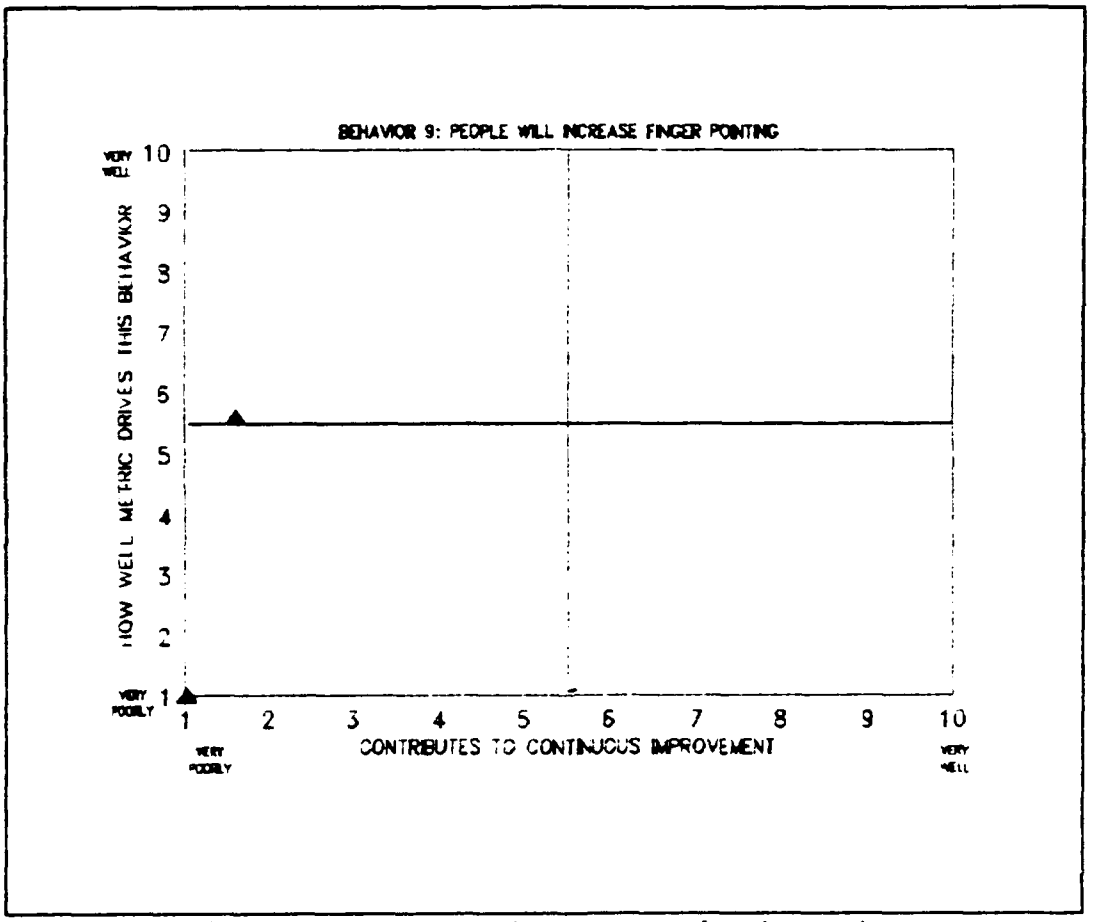


Figure A-9. Metric One, Behavior Nine

Behavior 10. People will seek to assure delivery process is reliable.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	-	3	-	-	7.00
How well does behavior contribute to CI?	-	-	-	-	-	1	1	1	1	1	8.00

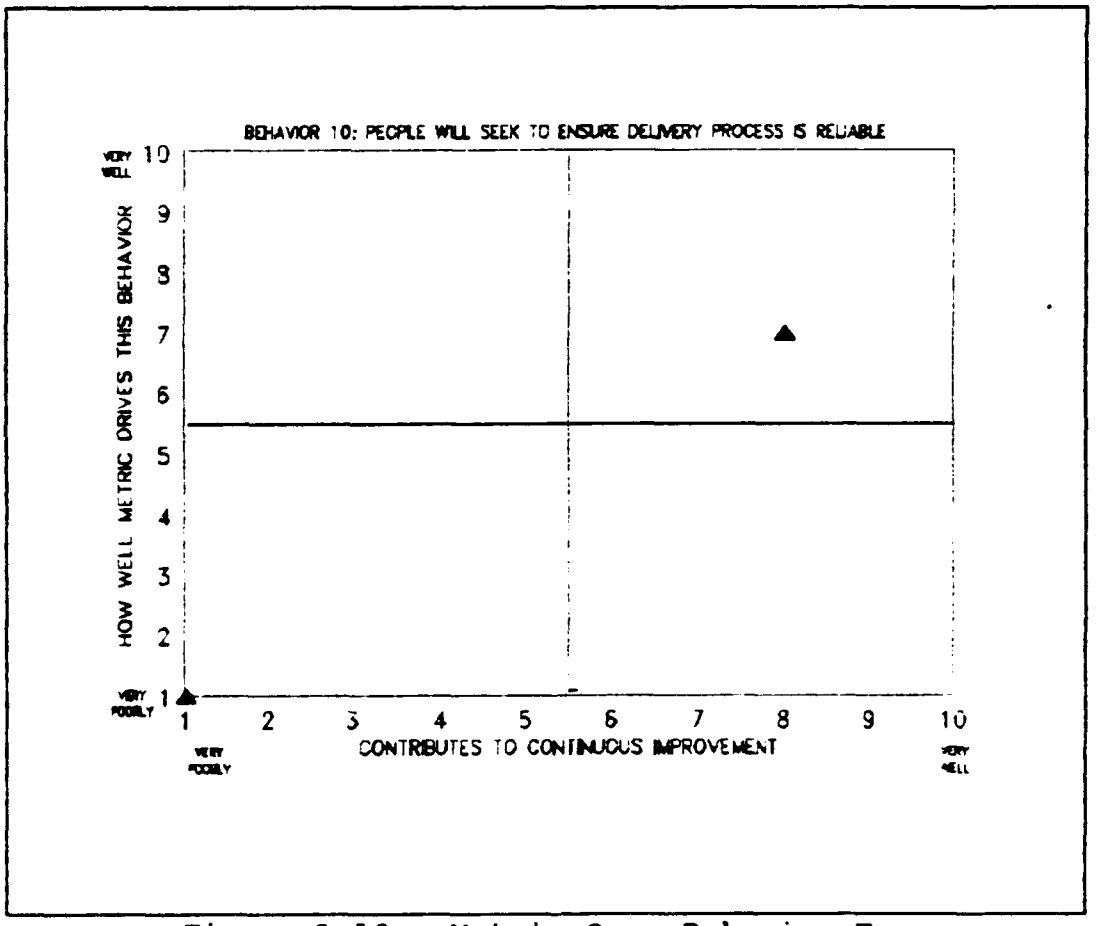


Figure A-10. Metric One, Behavior Ten

Behavior 11. People will concentrate on delivering exact number required.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	-	1	1	1	-	6.60
How well does behavior contribute to CI?	1	-	-	1	-	1	-	2	-	-	5.40

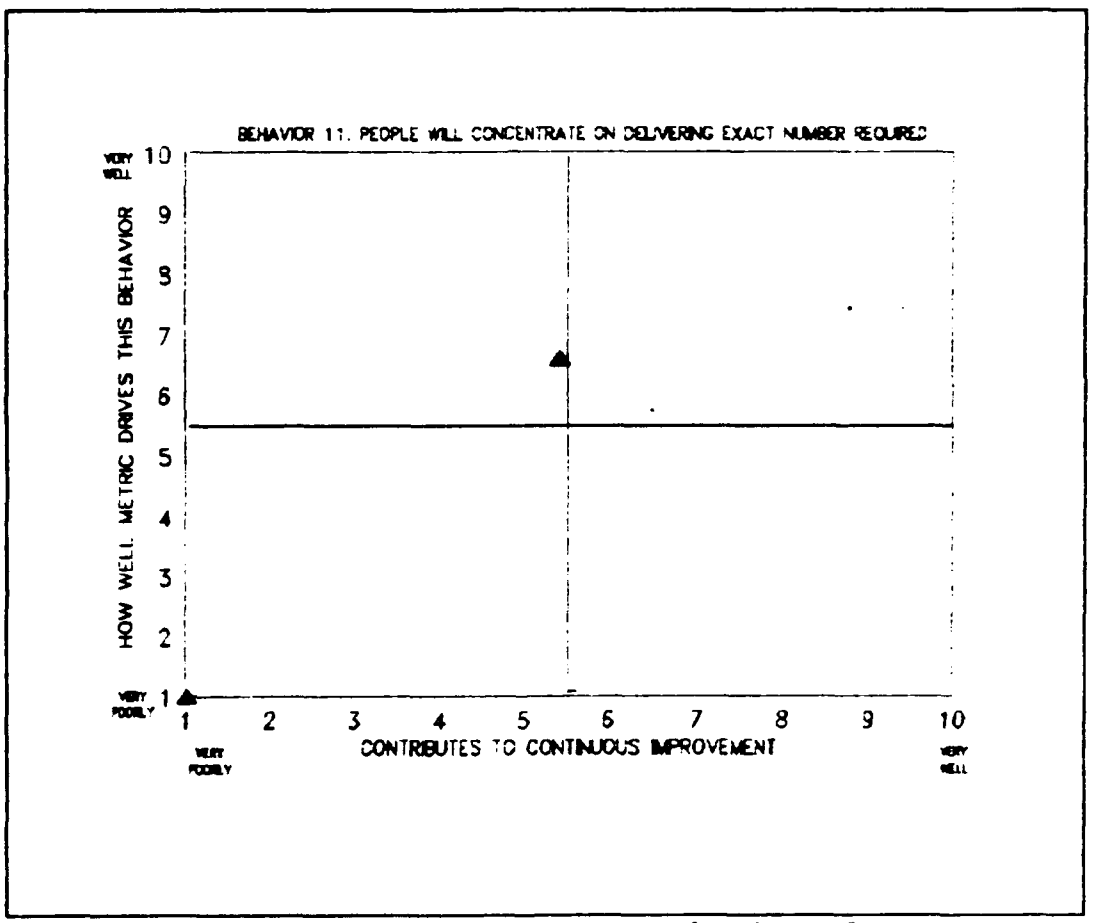


Figure A-11. Metric One, Behavior Eleven

Behavior 12. People will become more concerned with quantity than quality.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	-	2	1	2	9.00
How well does behavior contribute to CI?	1	2	-	1	-	1	-	-	-	-	3.00

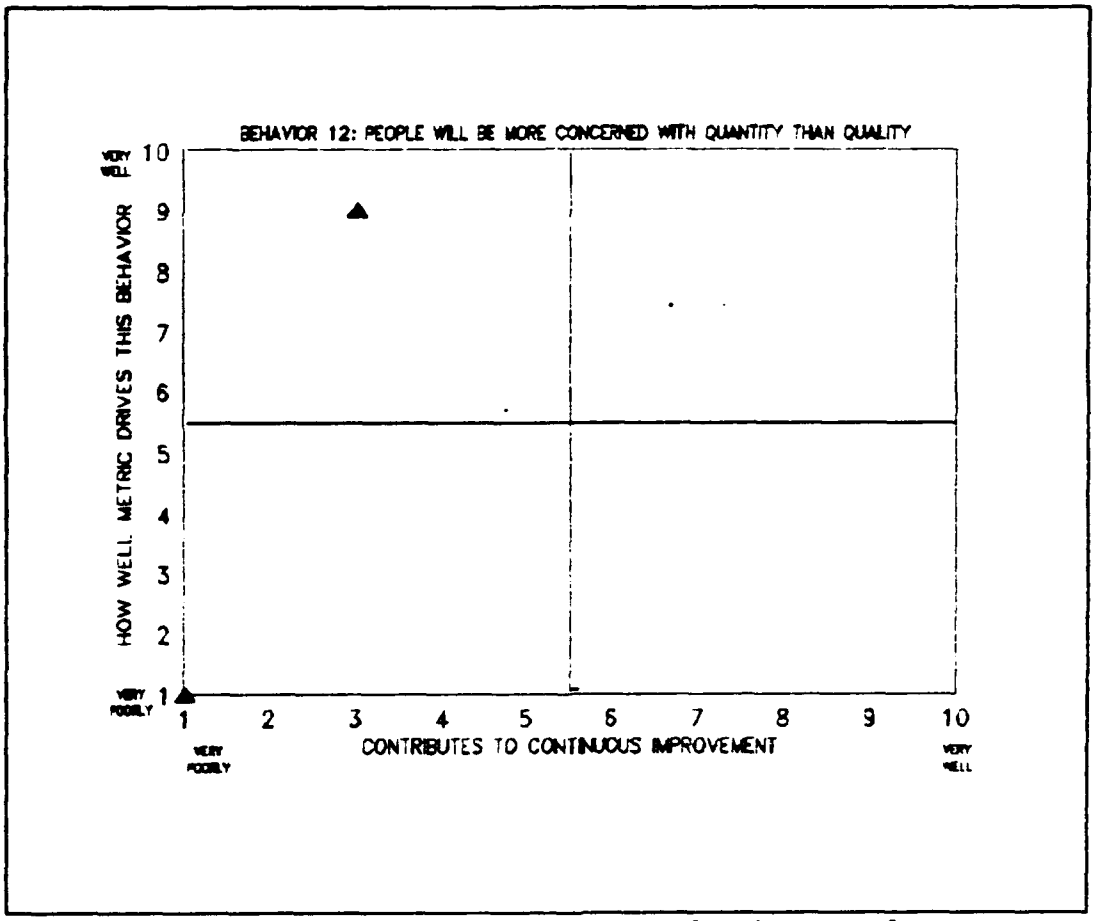


Figure A-12. Metric One, Behavior Twelve

Behavior 13. Increased emphasis on component availability during assembly process

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	1	-	1	7.20
How well does behavior contribute to CI?	-	-	-	-	-	-	3	-	-	2	8.20

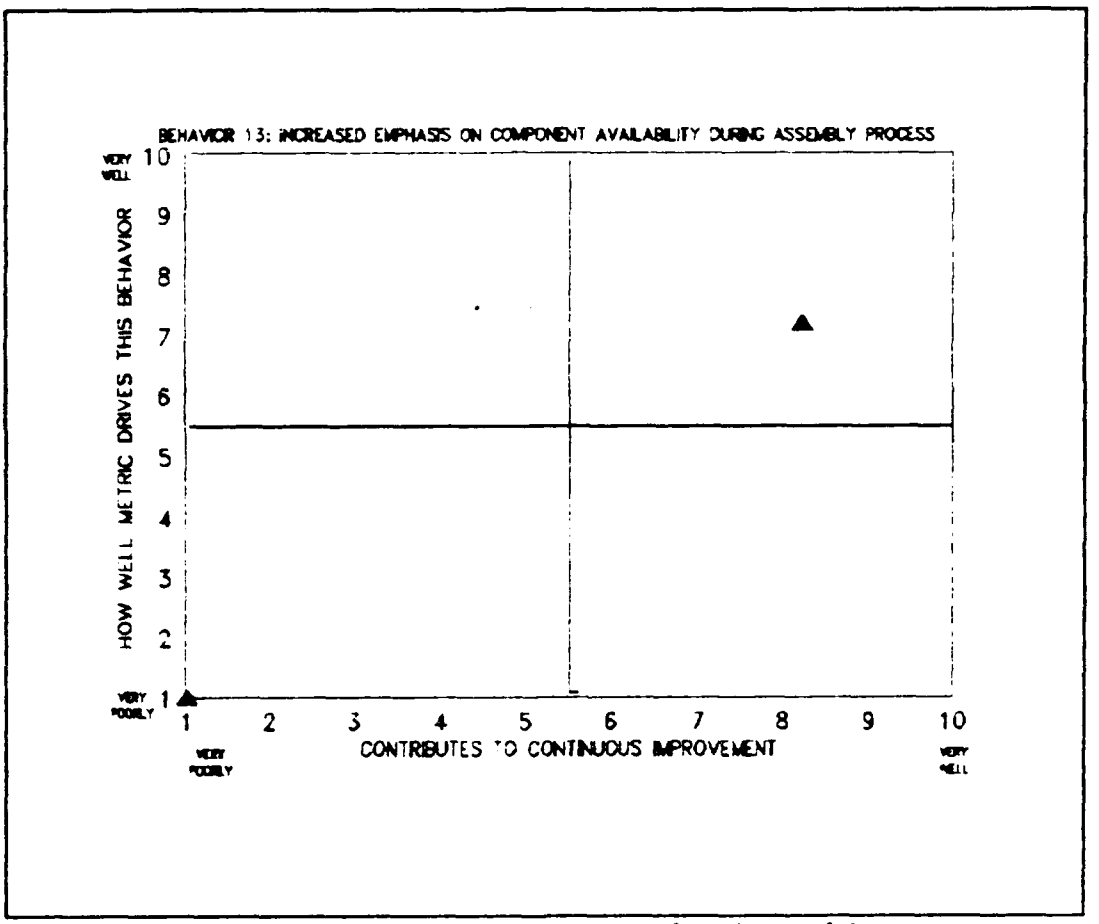


Figure A-13. Metric One, Behavior Thirteen

Behavior 14. People will put pressure on customers to accept product.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	2	2	-	-	7.00
How well does behavior contribute to CI?	3	1	-	-	-	-	1	-	-	-	2.40

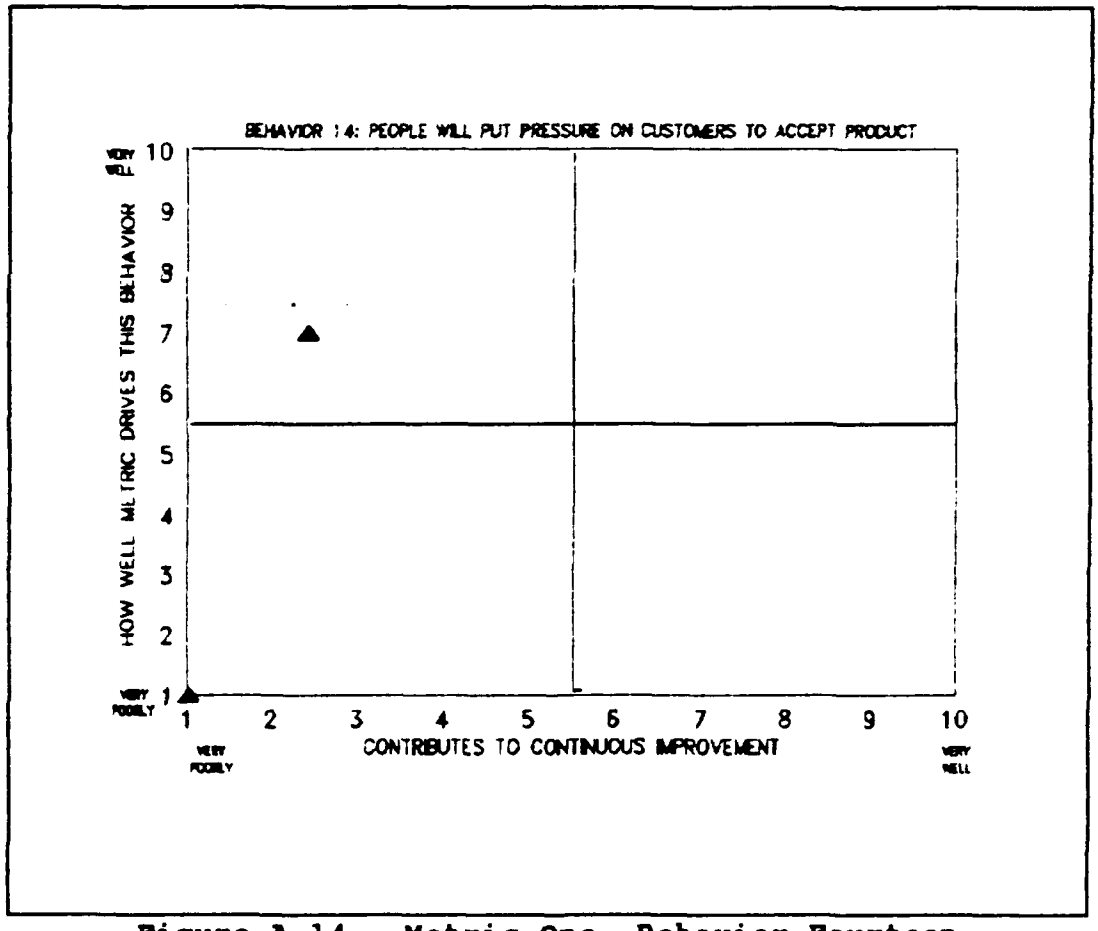


Figure A-14. Metric One, Behavior Fourteen

Behavior 15. Increased emphasis on statistical process control.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	2	-	-	1	-	6.00
How well does behavior contribute to CI?	-	-	-	-	-	-	2	-	1	2	8.60

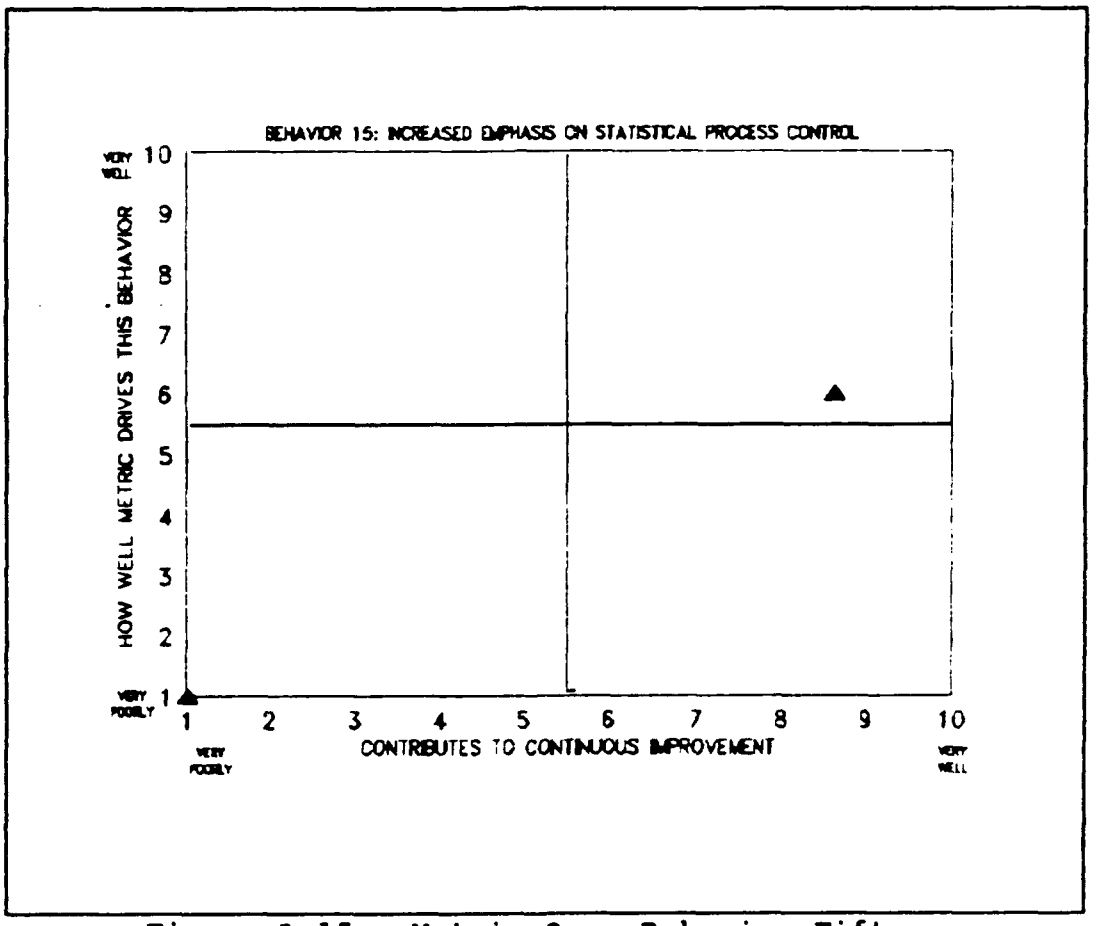


Figure A-15. Metric One, Behavior Fifteen

Metric 2 Behavior 1. Contractor will develop a process to reduce response time

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	2	1	-	2	8.40
How well does behavior contribute to CI?	-	-	-	-	-	-	1	1	2	1	8.60

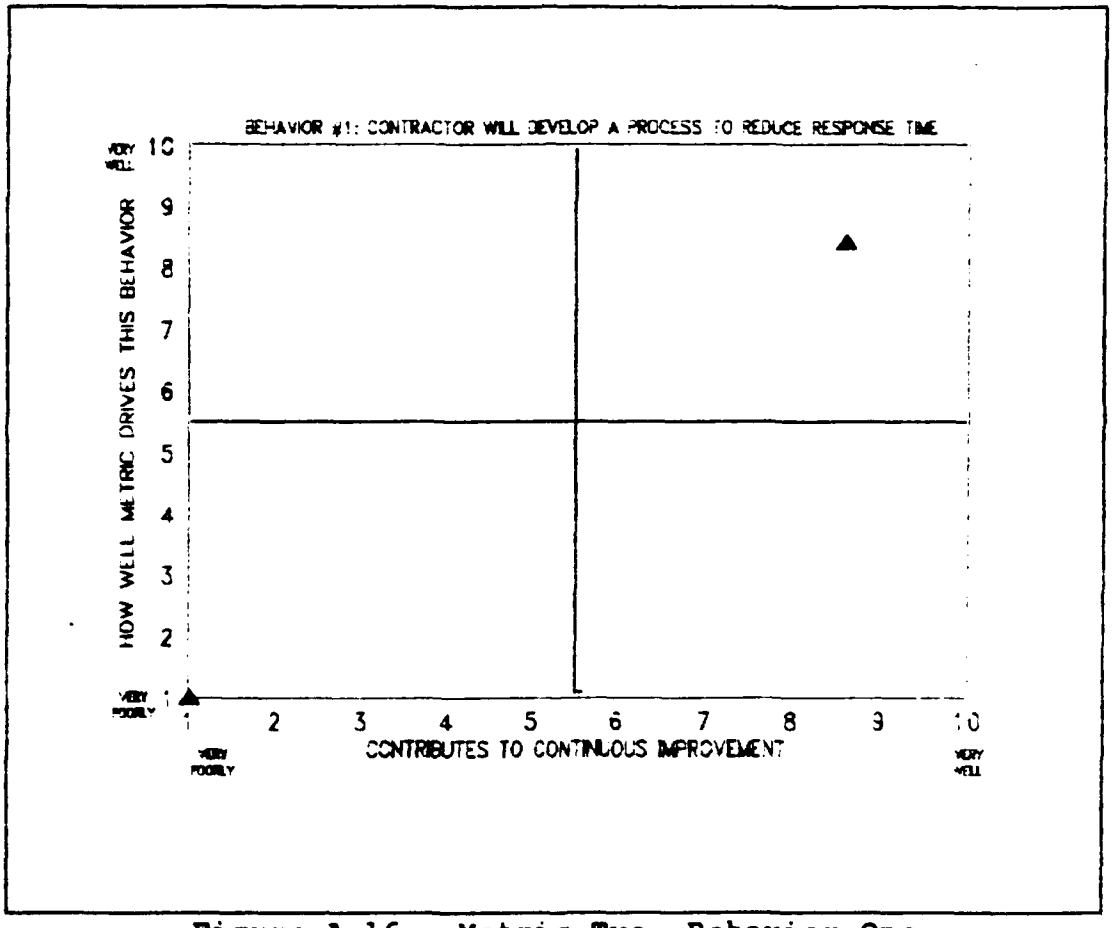


Figure A-16. Metric Two, Behavior One

Behavior 2. Contractor emphasis on response time instead of response content

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	2	3	-	-	7.60
How well does behavior contribute to CI?	1	2	-	2	-	-	-	-	-	-	2.60

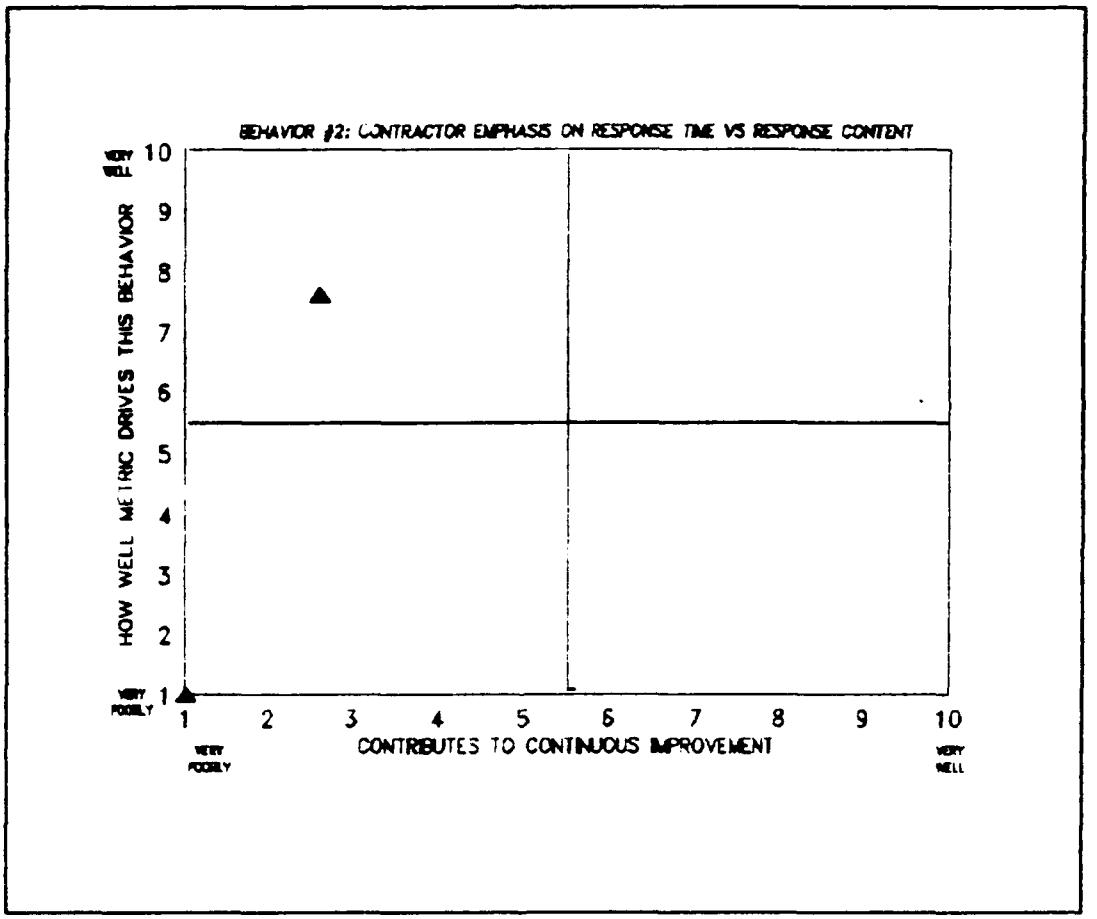


Figure A-17. Metric Two, Behavior Two

Behavior 3. Contractor shifts manpower to quickly work the SRs

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	-	1	3	-	8.00
How well does behavior contribute to CI?	1	1	1	-	-	-	1	1	-	-	4.20

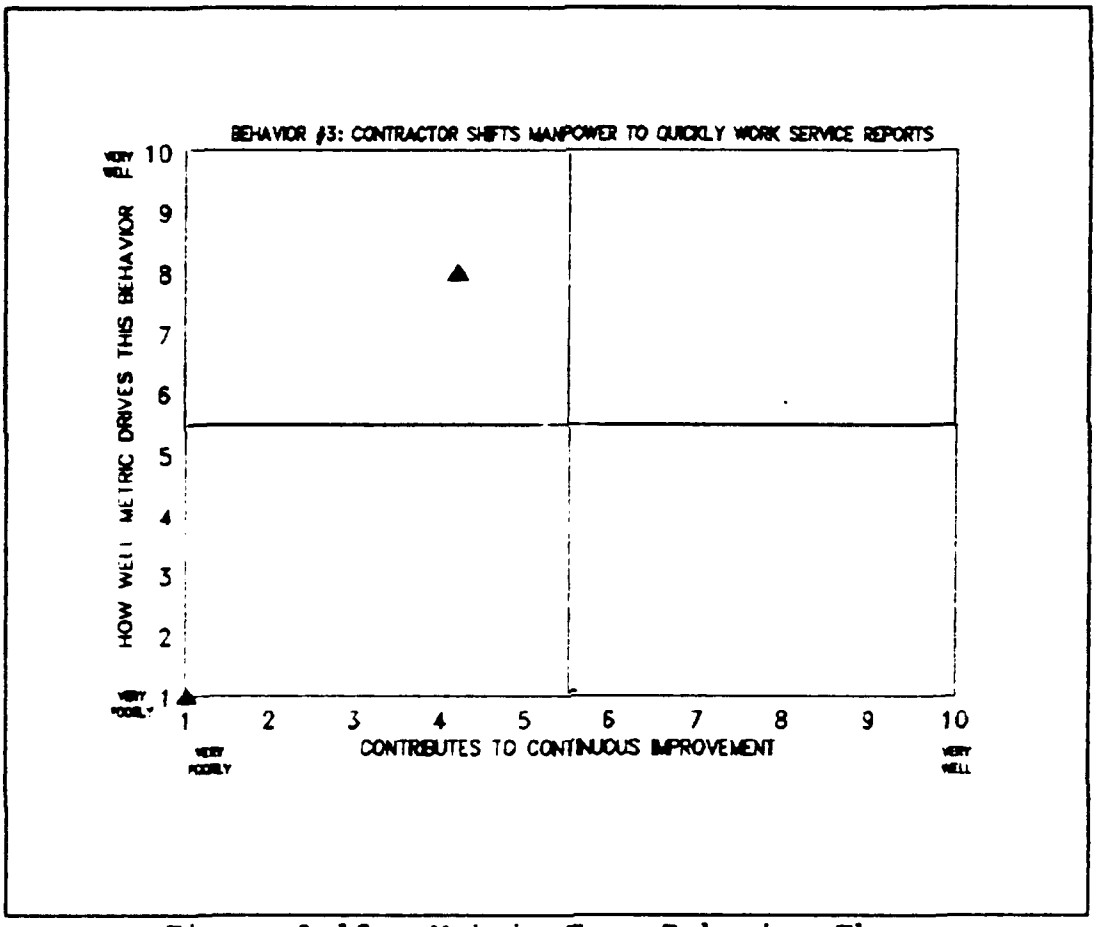


Figure A-18. Metric Two, Behavior Three

Behavior 4. Increased SPO service report status reviews

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric	-	-	-	1	-	2	-	1	-	1	6.80
How well does behavior contribute to CI?	-	-	2	-	-	1	-	1	-	1	6.00

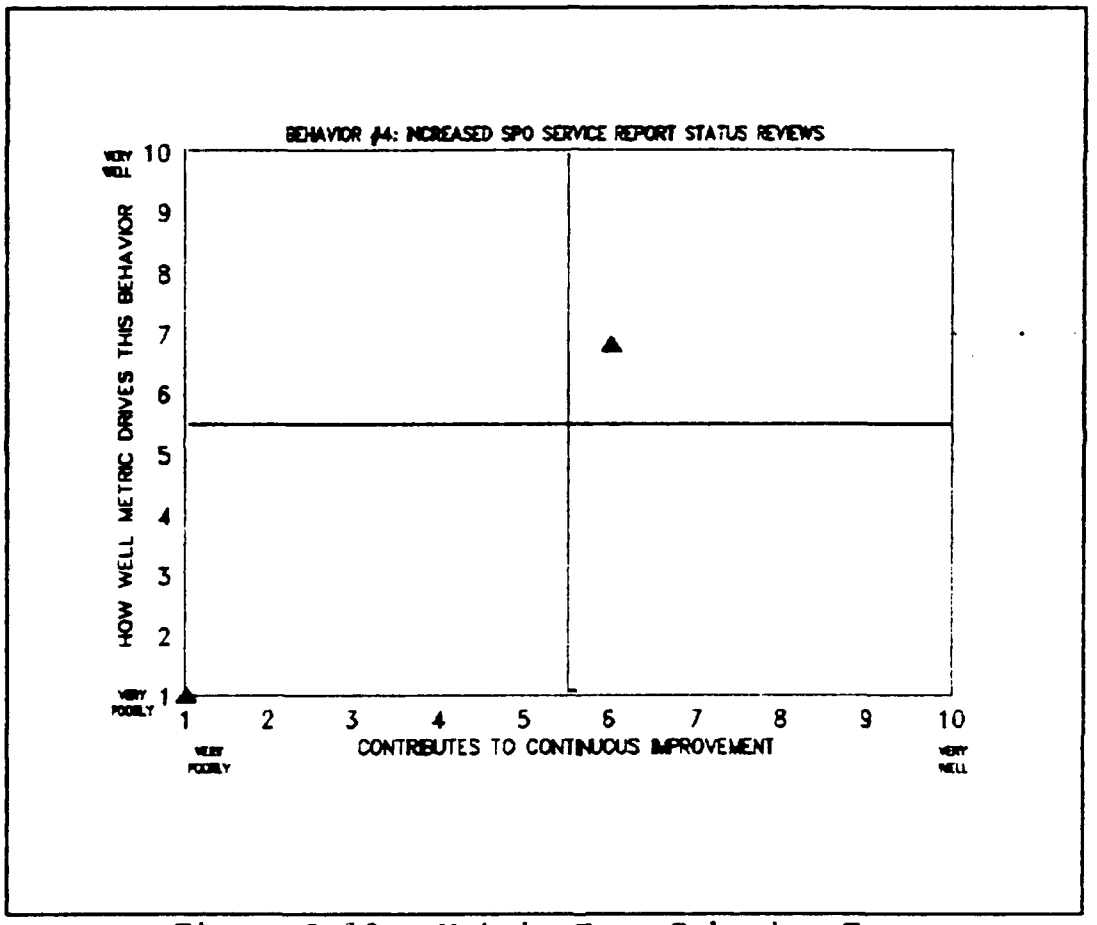


Figure A-19. Metric Two, Behavior Four

Behavior 5. SPO will pressure contractor to respond faster

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	1	1	1	1	8.00
How well does behavior contribute to CI?	-	1	1	2	-	-	1	-	-	-	4.00

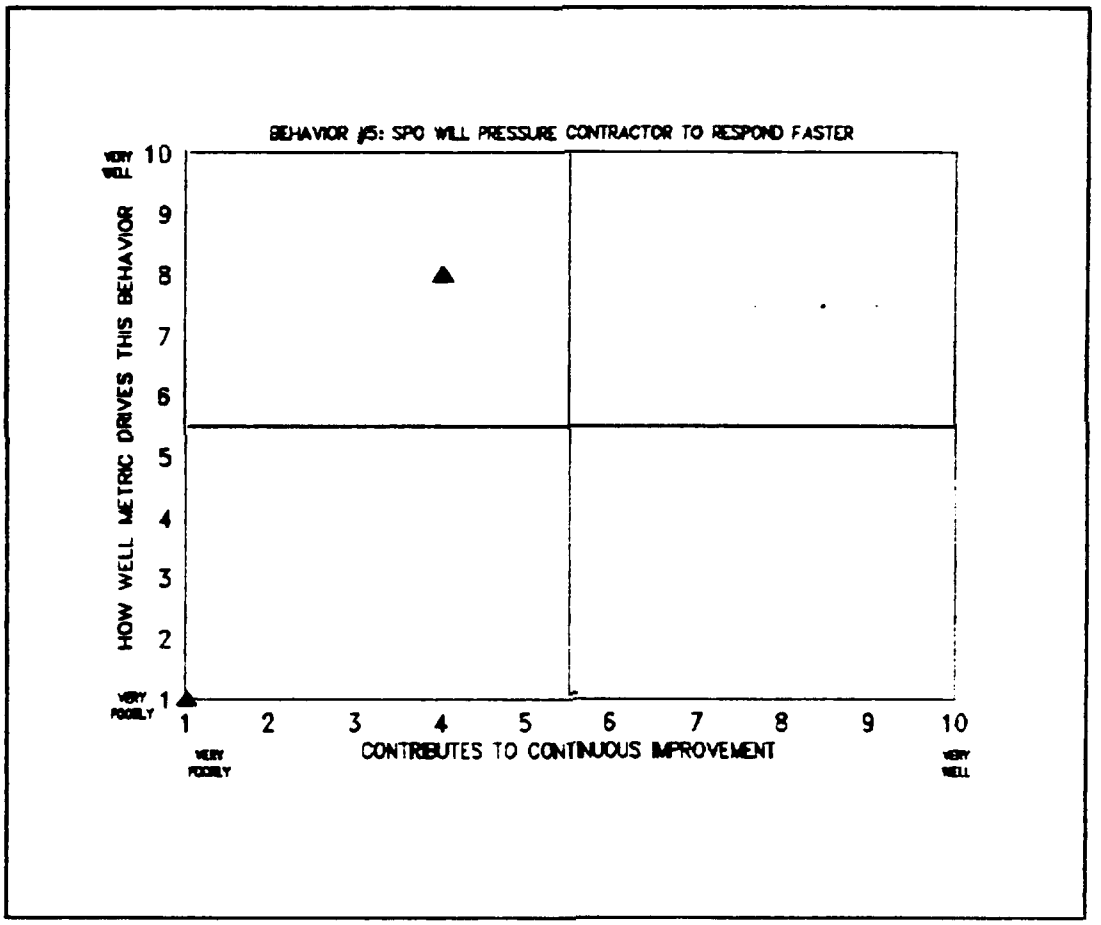


Figure A-20. Metric Two, Behavior Five

Behavior 6. SPO will develop a process to resolve a higher rate of SRs

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	1	1	2	-	-	-	1	-	-	4.20
How well does behavior contribute to CI?	-	-	-	1	-	-	2	-	1	1	7.40

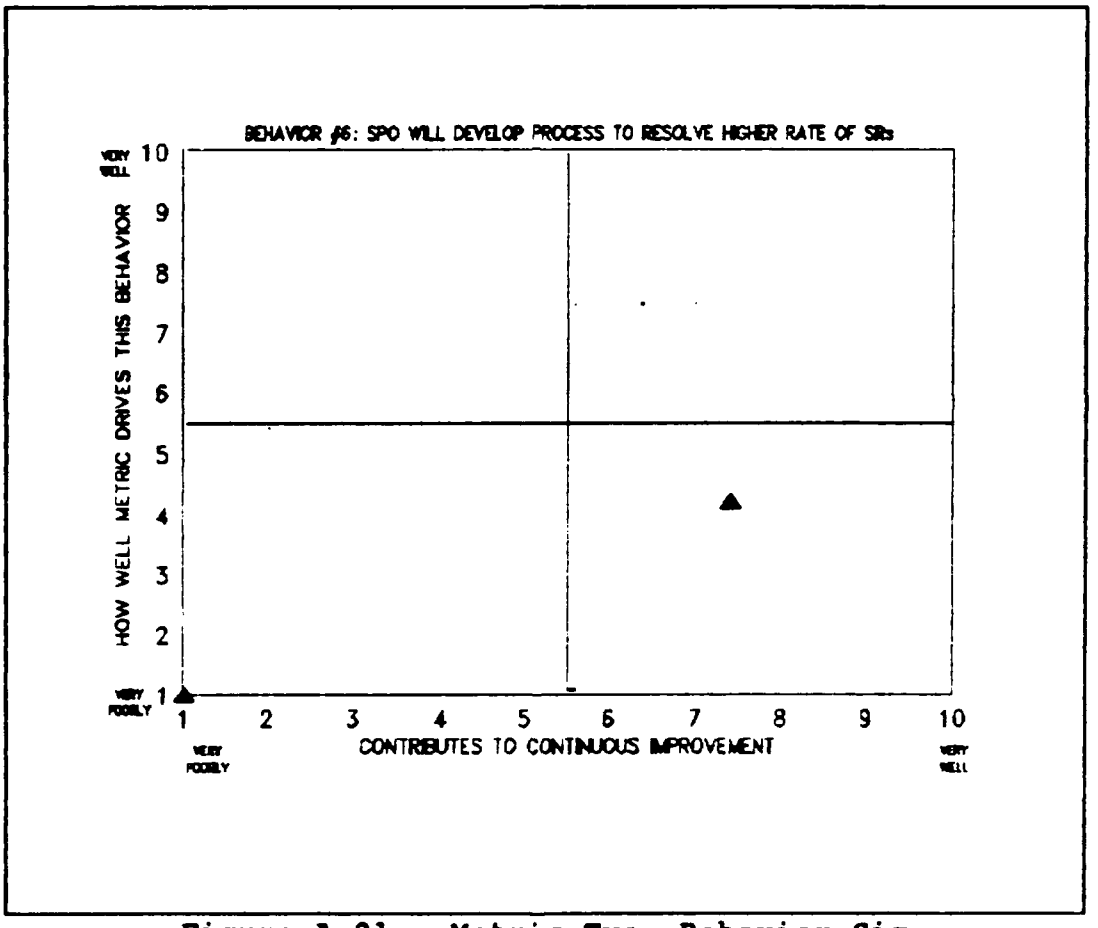


Figure A-21. Metric Two, Behavior Six

Behavior 7. SPO emphasis on contractor response time instead of SPO SR resolution

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	2	1	1	-	7.40
How well does behavior contribute to CI?	-	3	1	-	-	-	1	-	-	-	3.20

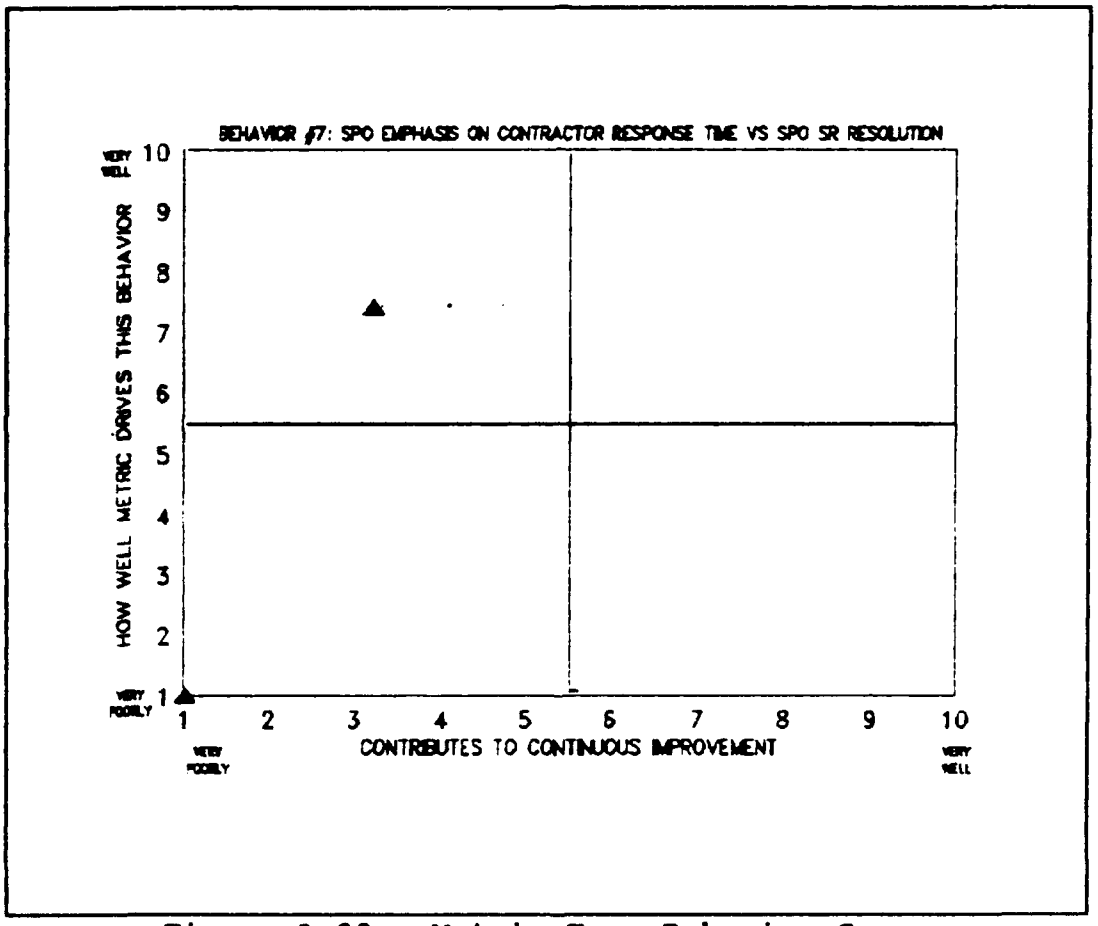


Figure A-22. Metric Two, Behavior Seven

Behavior 8. Contractor tries to get heads-up before problems are officially submitted

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	1	-	1	7.20
How well does behavior contribute to CI?	-	-	-	-	-	-	1	1	2	1	8.60

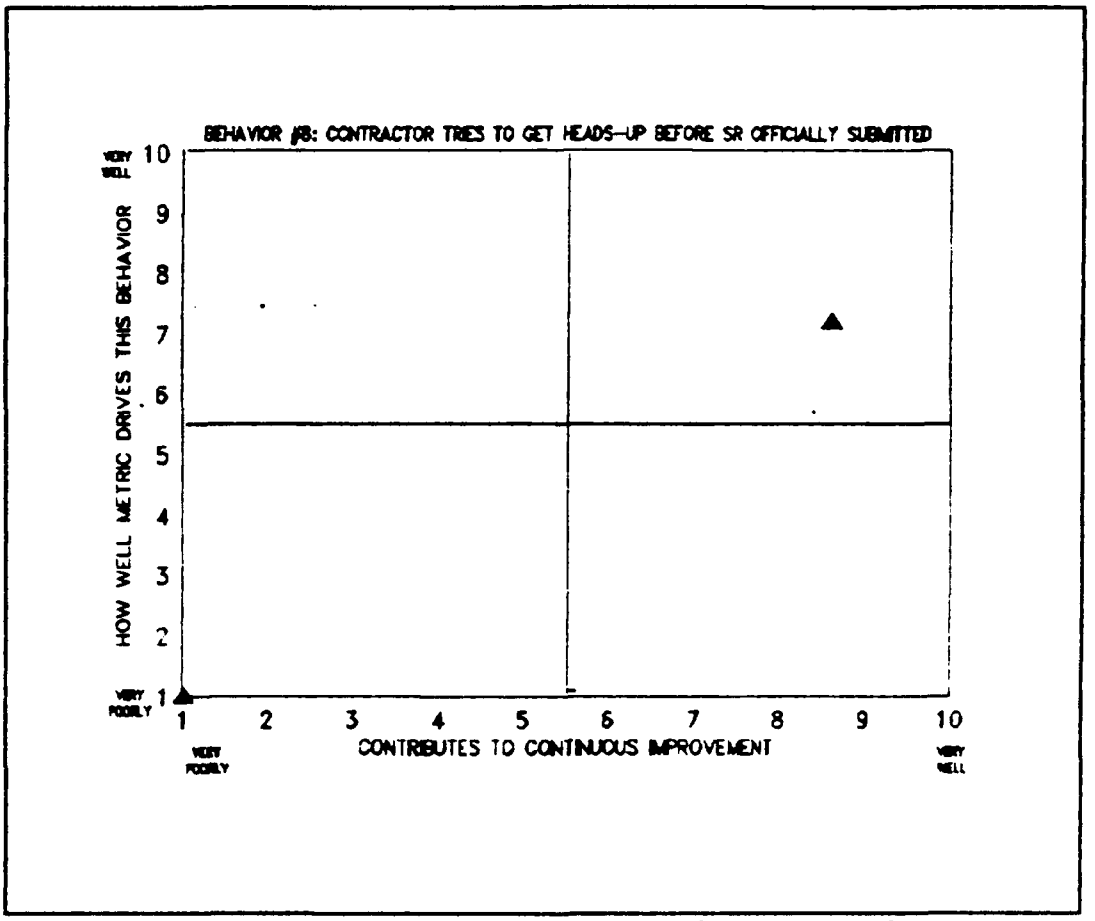


Figure A-23. Metric Two, Behavior Eight

Behavior 9. Customer will feel his SRs are important; may increase SRs

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	1	-	2	-	-	6.20
How well does behavior contribute to CI?	1	-	1	1	-	-	-	1	1	-	5.00

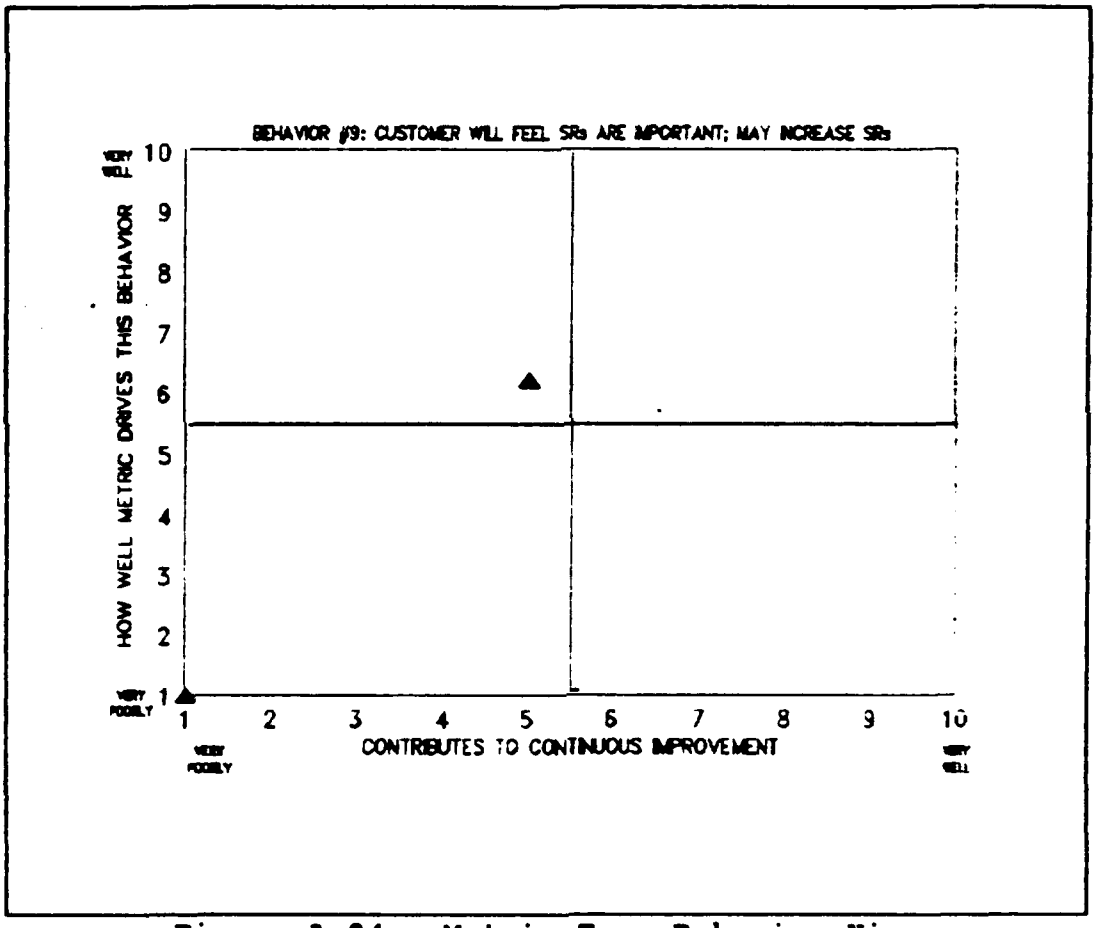


Figure A-24. Metric Two, Behavior Nine

Metric 3 Behavior 1. Increased SPO contracting reviews for UCA activties

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	1	1	7.40
How well does behavior contribute to CI?	-	-	-	1	-	2	-	1	-	1	6.80

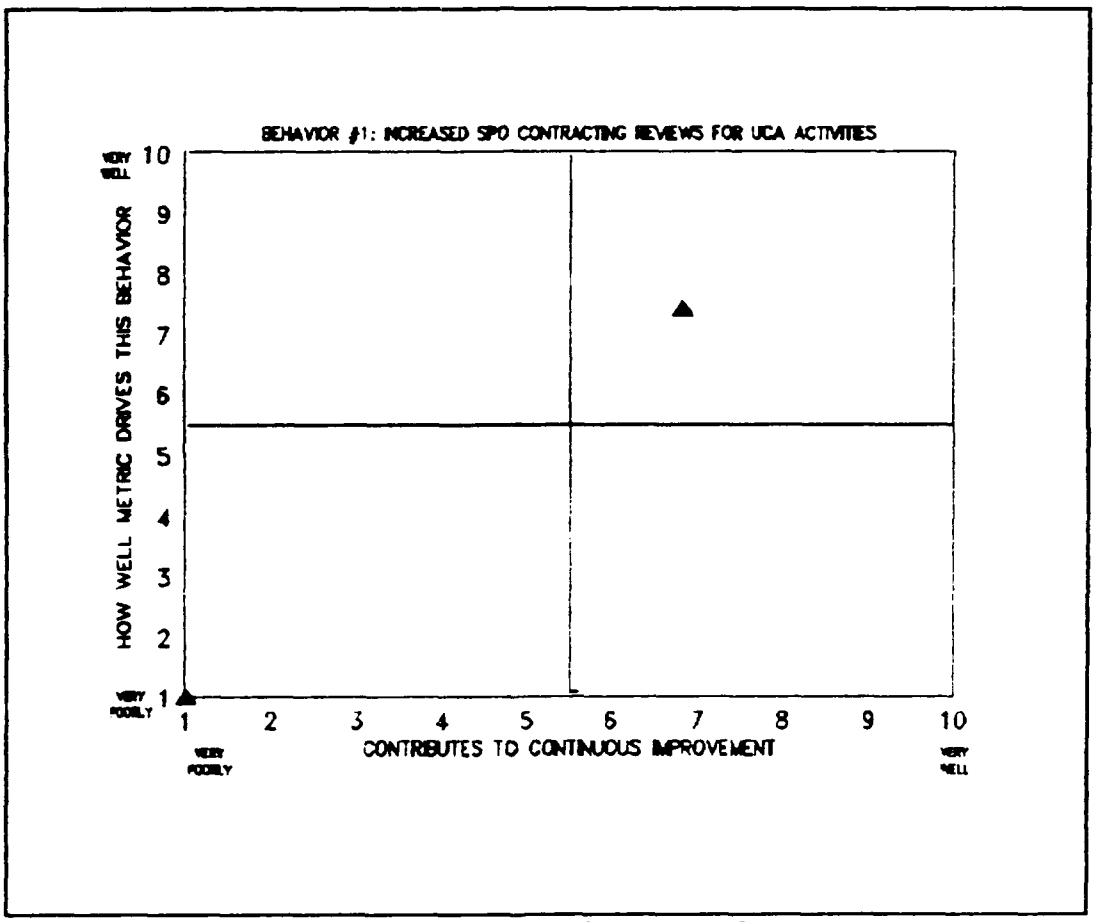


Figure A-25. Metric Three, Behavior One

Behavior 2. SPO will decrease detail of technical evaluation

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	-	2	1	-	7.20
How well does behavior contribute to CI?	1	1	1	1	-	1	-	-	-	-	3.20

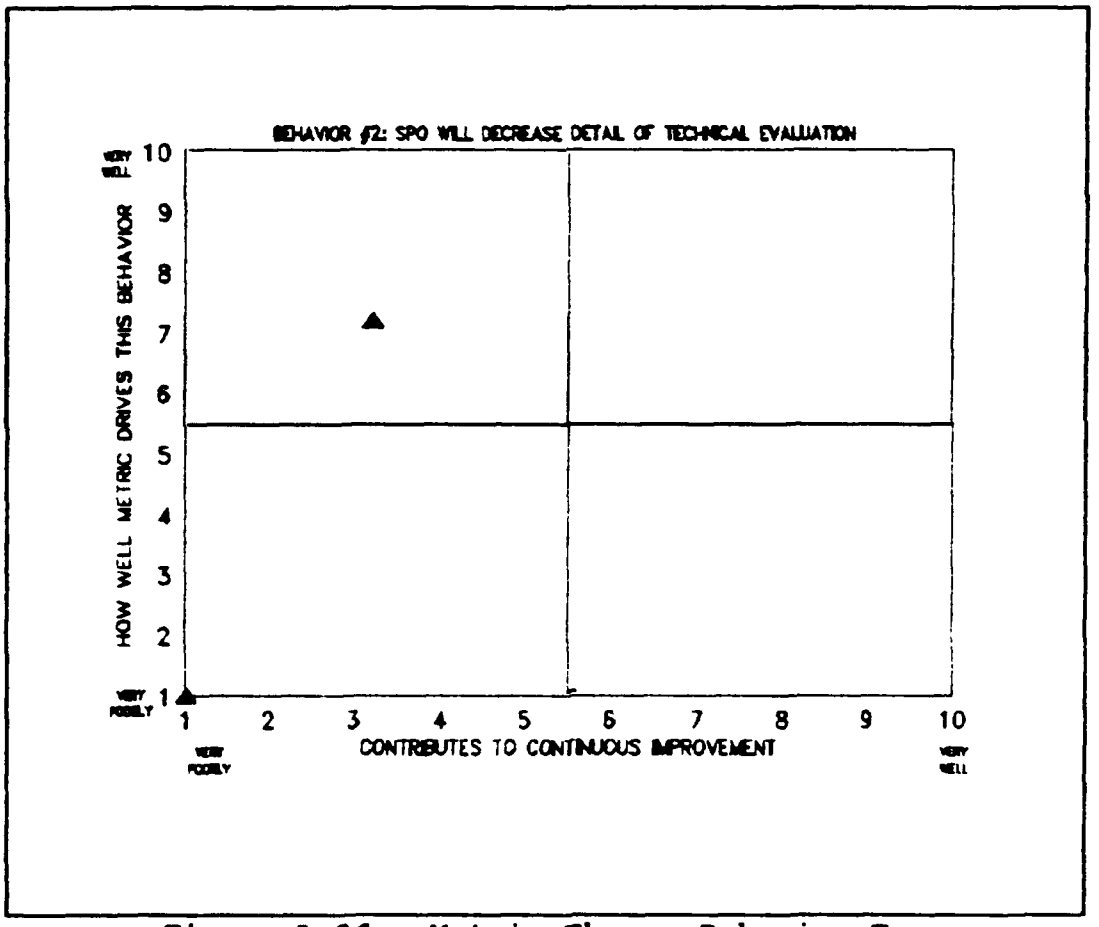


Figure A-26. Metric Three, Behavior Two

Behavior 3. Increased emphasis on SPO UCA definitization process

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric	-	-	-	-	-	-	-	4	-	1	8.40
How well does behavior contribute to CI?	-	-	-	-	-	-	1	2	-	2	8.60

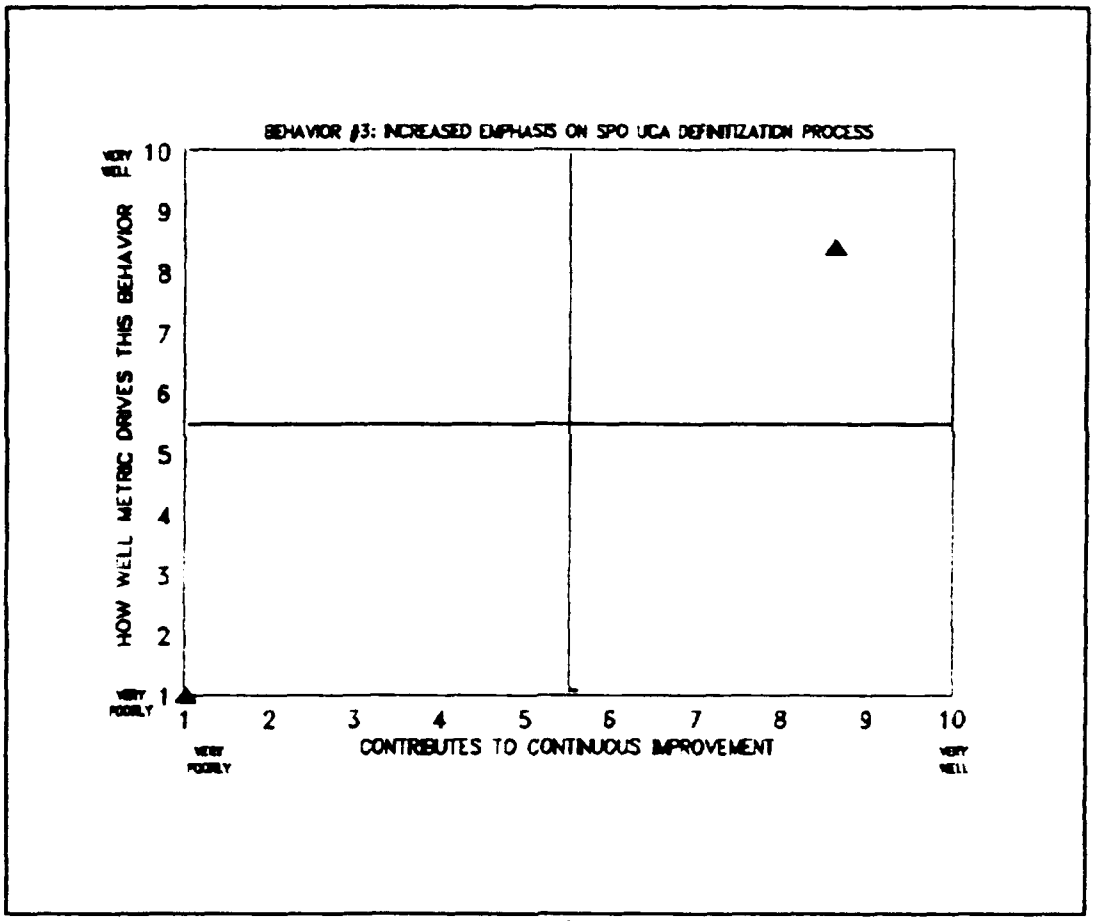


Figure A-27. Metric Three, Behavior Three

Behavior 4. SPO emphasis on striking ANY deal when approaching the 180 day limit

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	-	-	-	3	1	-	7.40
How well does behavior contribute to CI?	1	2	1	-	-	-	-	1	-	-	3.20

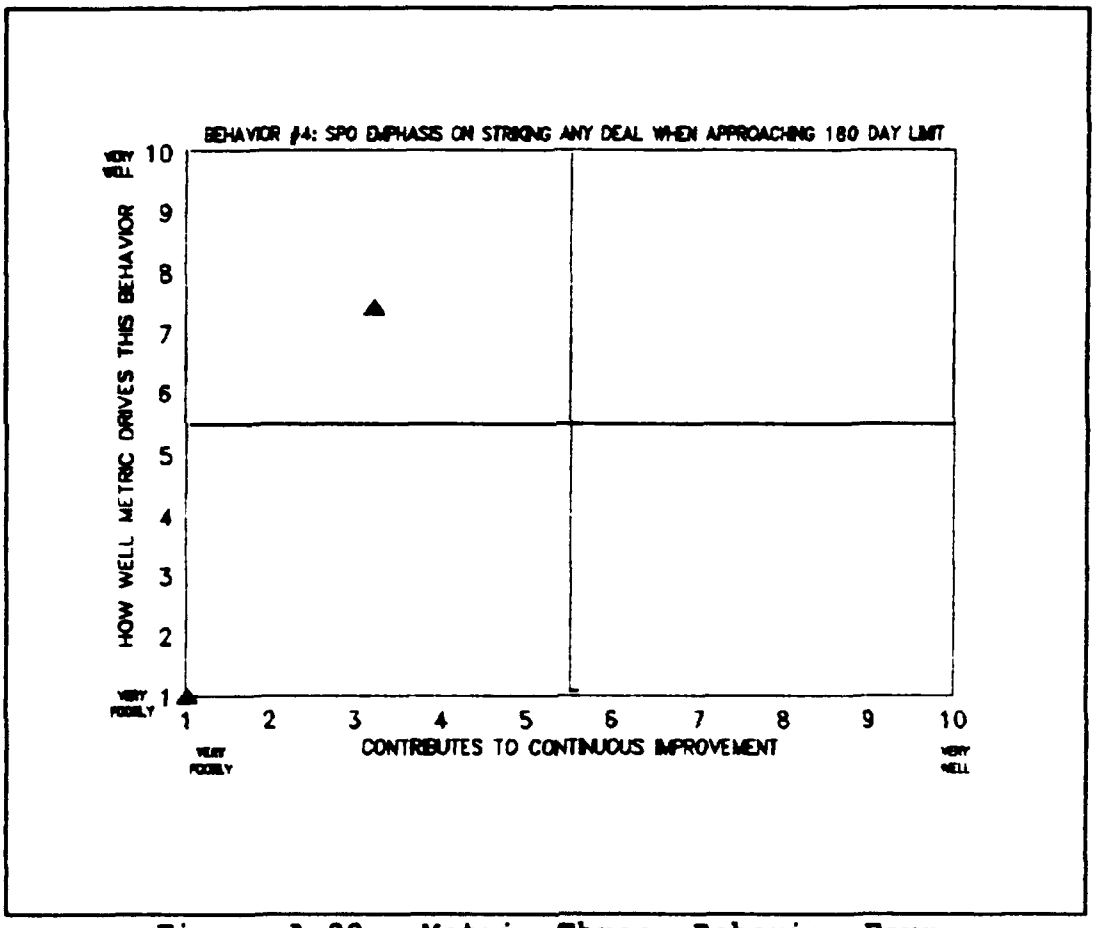


Figure A-28. Metric Three, Behavior Four

Behavior 5. Lots of activity as the 180 day limit approaches

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	-	2	1	1	8.20
How well does behavior contribute to CI?	-	-	2	-	2	-	-	1	-	-	4.80

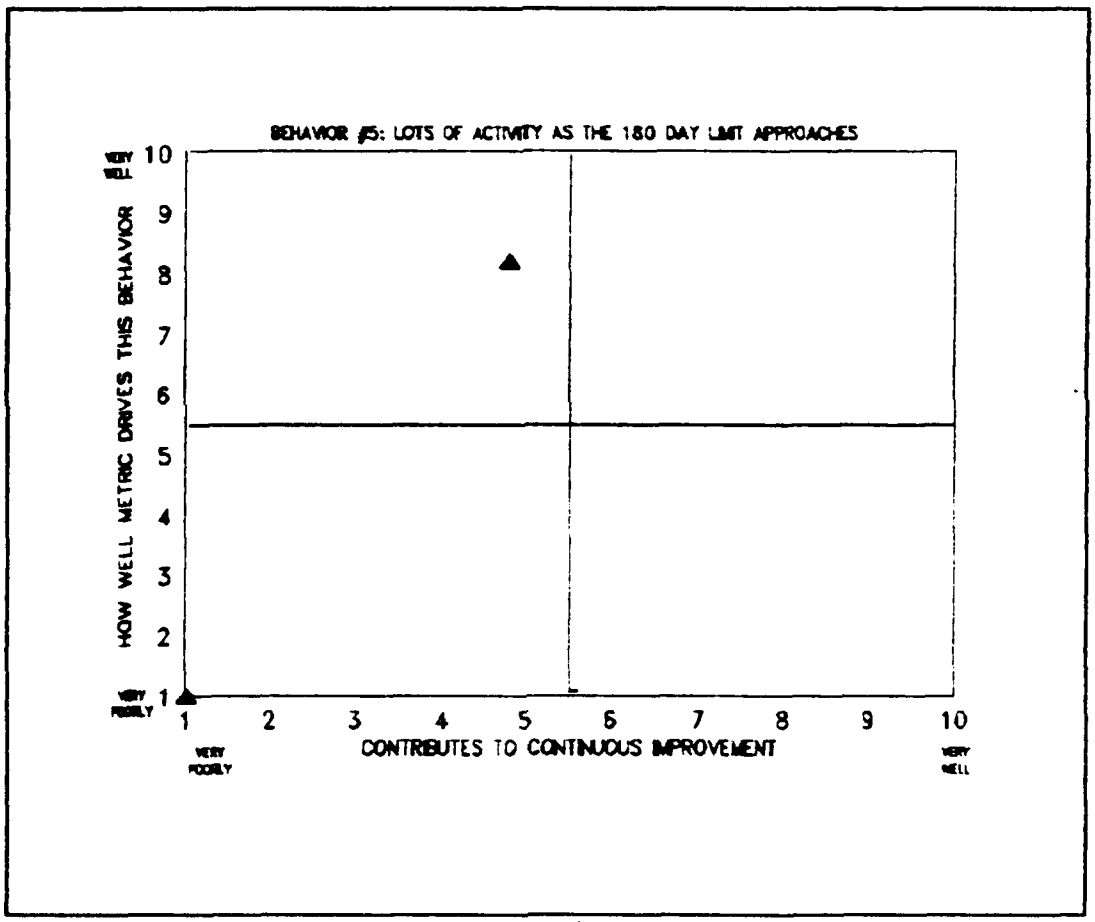


Figure A-29. Metric Three, Behavior Five

Behavior 6. SPO negotiates faster with less concern for price when approaching 180 day limit

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	-	2	2	-	8.00
How well does behavior contribute to CI?	1	1	2	-	-	-	-	1	-	-	3.40

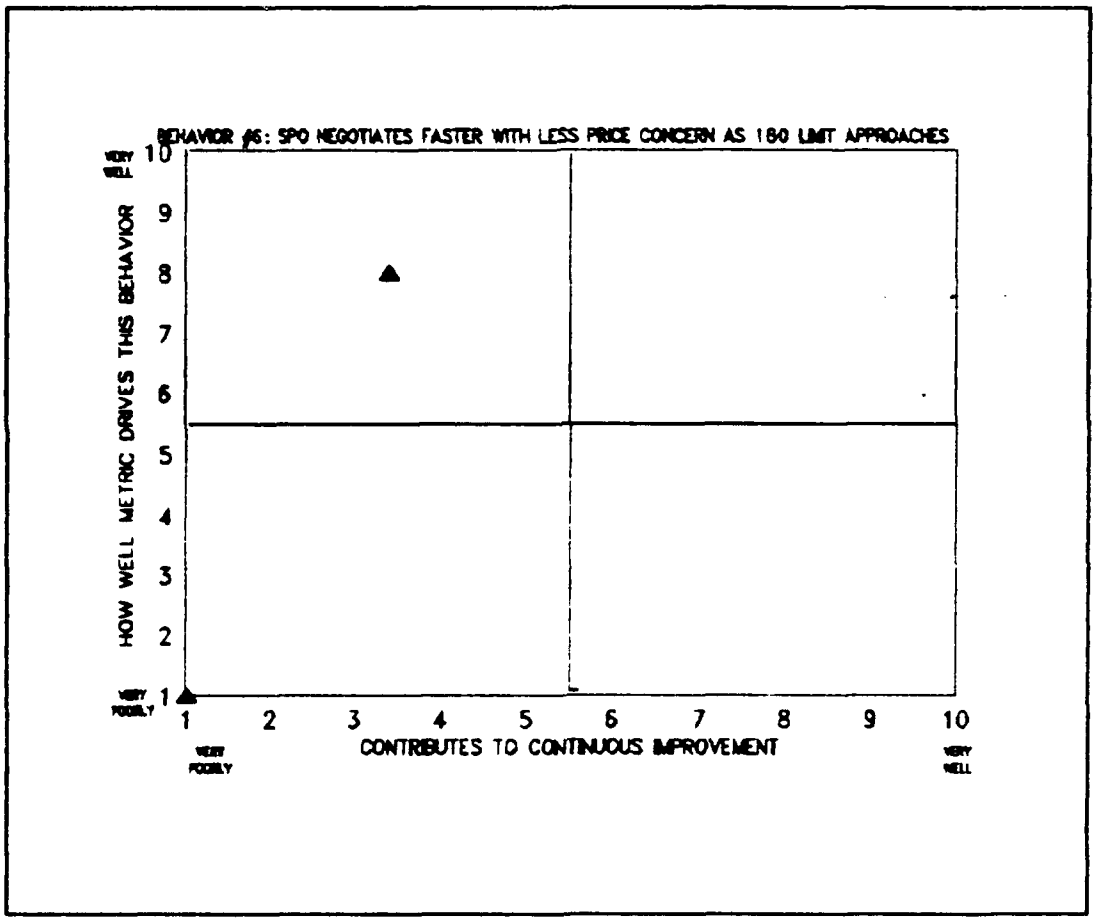


Figure A-30. Metric Three, Behavior Six

Behavior 7. Contractor will slow down negotiating process to corner Gov't

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	2	-	-	2	-	1	7.20
How well does behavior contribute to CI?	1	2	1	1	-	-	-	-	-	-	2.40

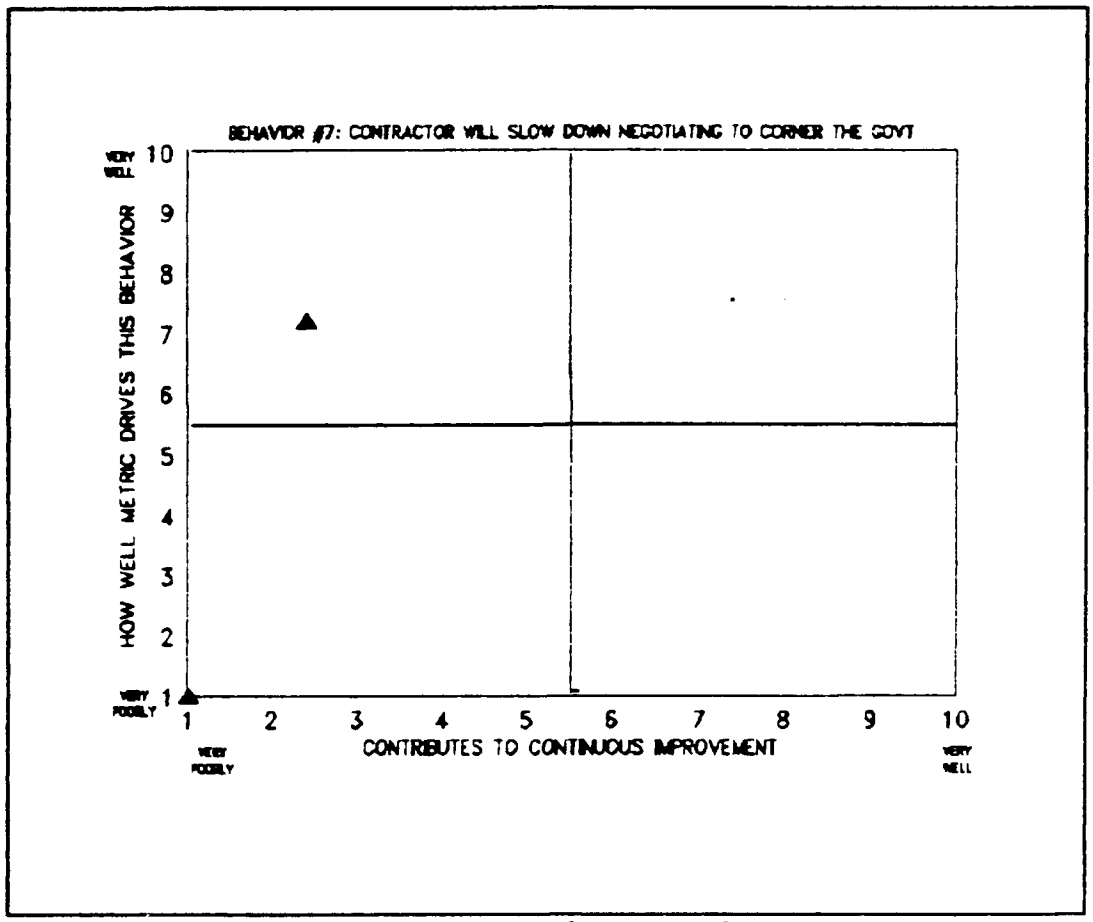


Figure A-31. Metric Three, Behavior Seven

Behavior 8. Increased emphasis on contractor UCA proposal development process

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	1	-	1	1	-	1	6.60
How well does behavior contribute to CI?	-	-	-	-	-	-	1	-	2	2	9.00

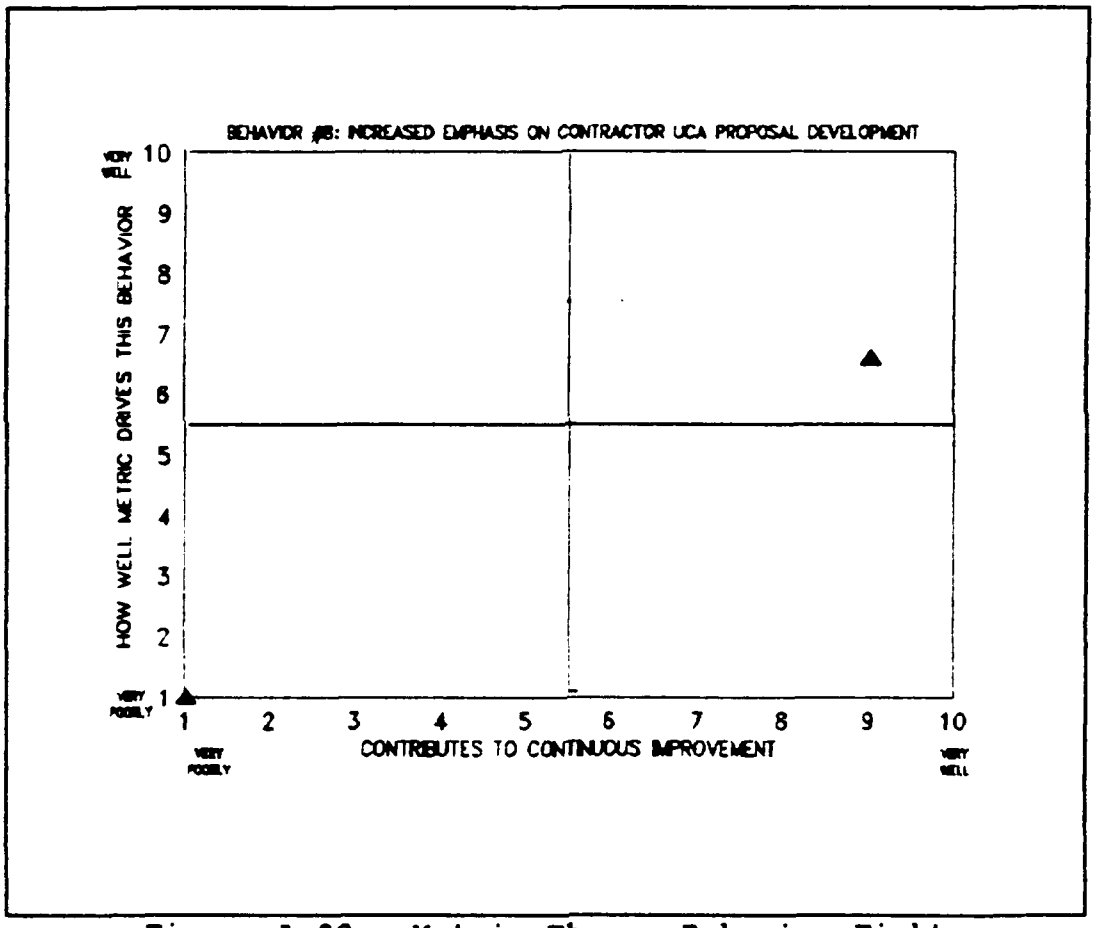


Figure A-32. Metric Three, Behavior Eight

Behavior 9. SPO UCA definitization priority established by UCA age

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	1	1	1	1	7.80
How well does behavior contribute to CI?	-	-	3	-	-	1	-	1	-	-	4.60

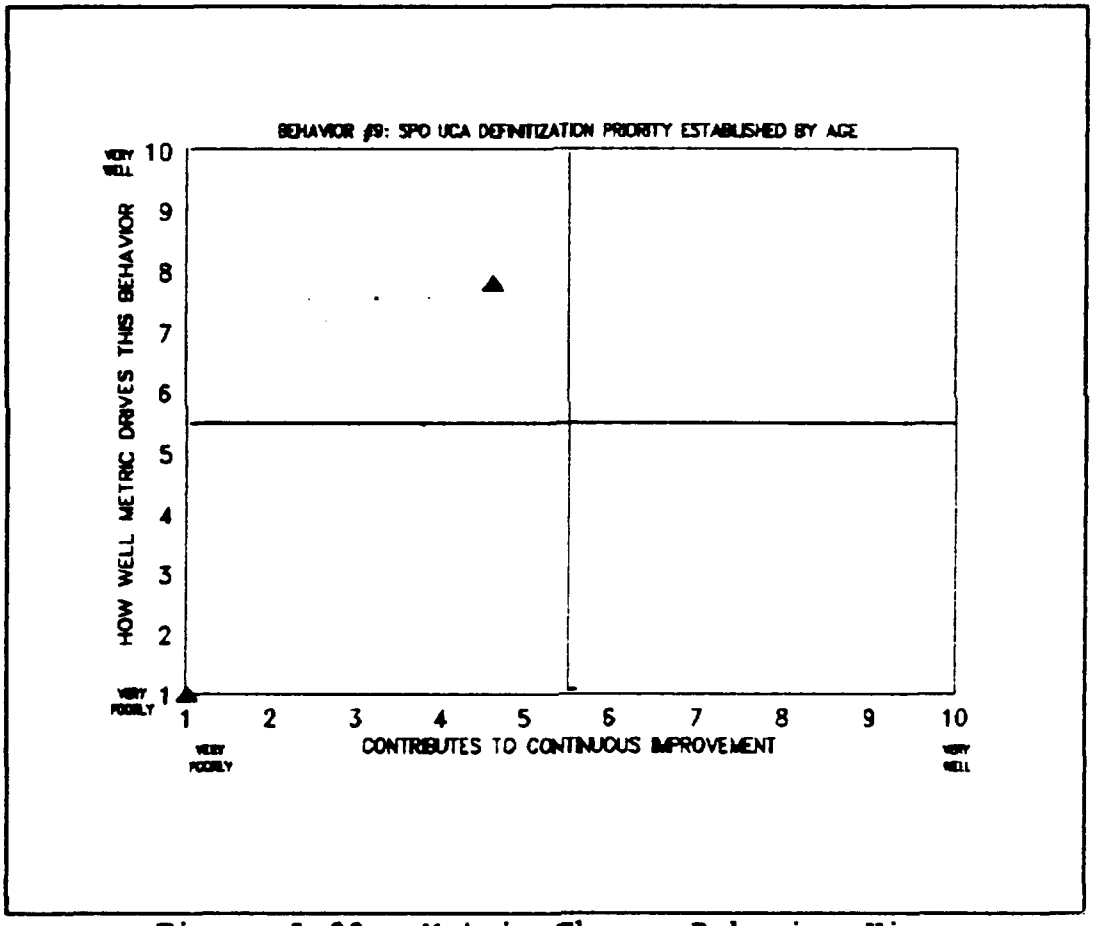


Figure A-33. Metric Three, Behavior Nine

Behavior 10. SPO will decrease use of UCAs unless hard core requirement really exists

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	2	-	7.20
How well does behavior contribute to CI?	-	1	-	-	-	-	-	2	1	1	7.40

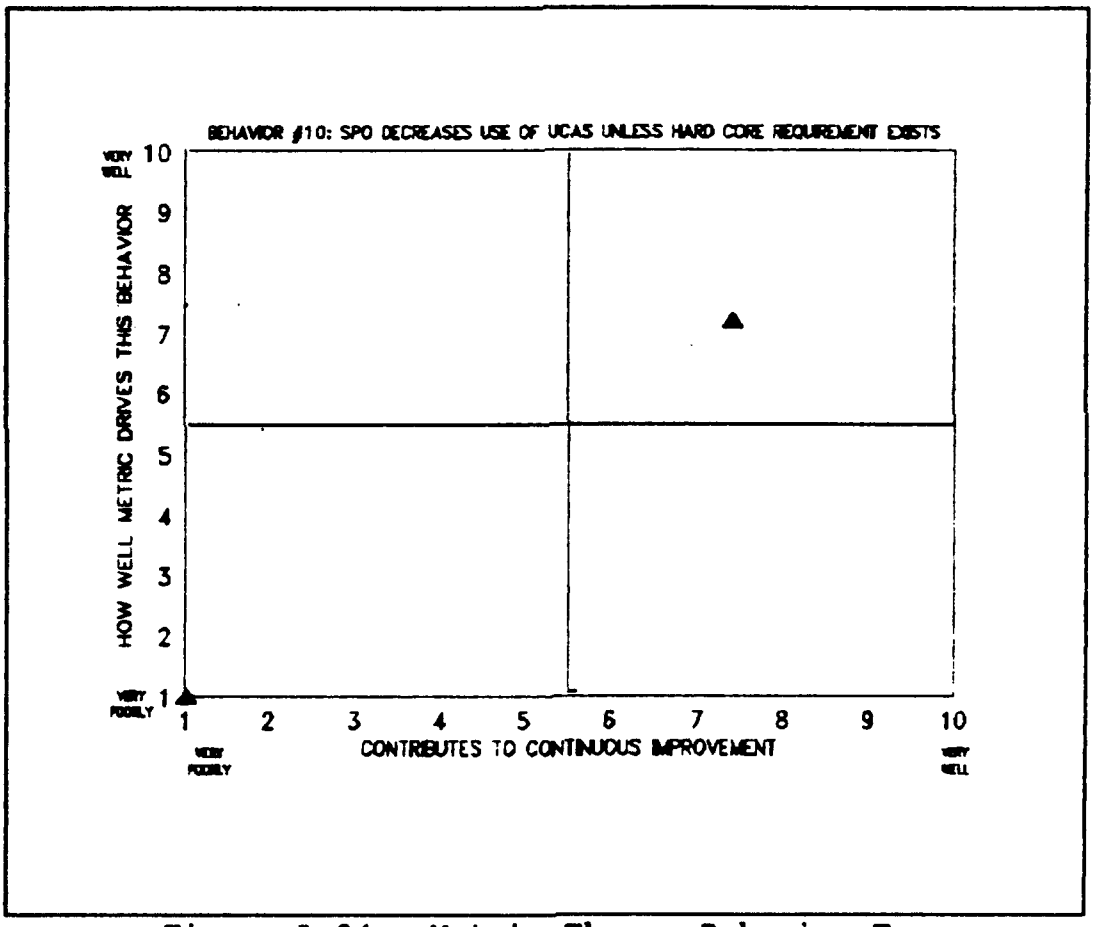


Figure A-34. Metric Three, Behavior Ten

Behavior 11. Decreased proposal evaluation periods

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	3	1	-	-	1	7.00
How well does behavior contribute to CI?	-	1	1	1	-	-	-	1	-	1	5.40

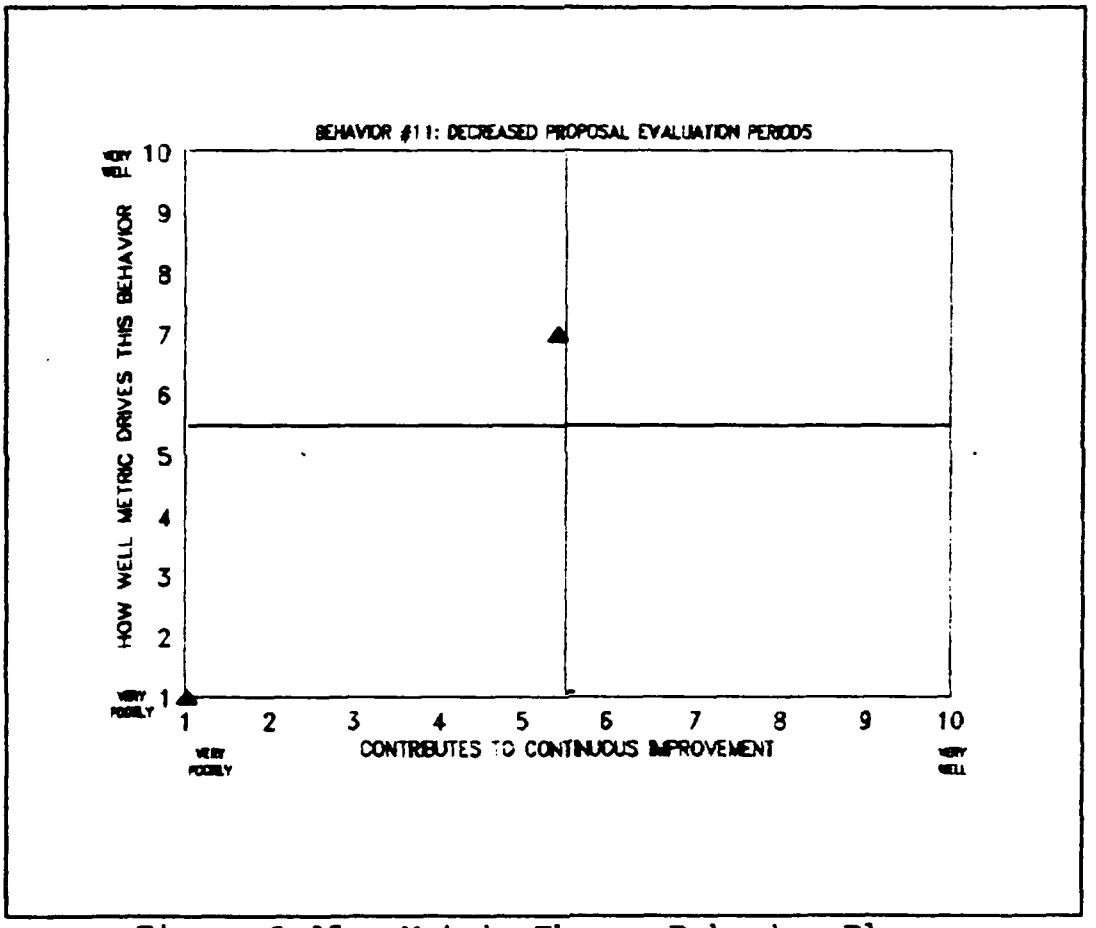


Figure A-35. Metric Three, Behavior Eleven

Behavior 12. Increased manpower working UCAs older than 180 days

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	-	1	-	1	1	7.00
How well does behavior contribute to CI?	-	-	1	1	-	1	-	-	2	-	6.20

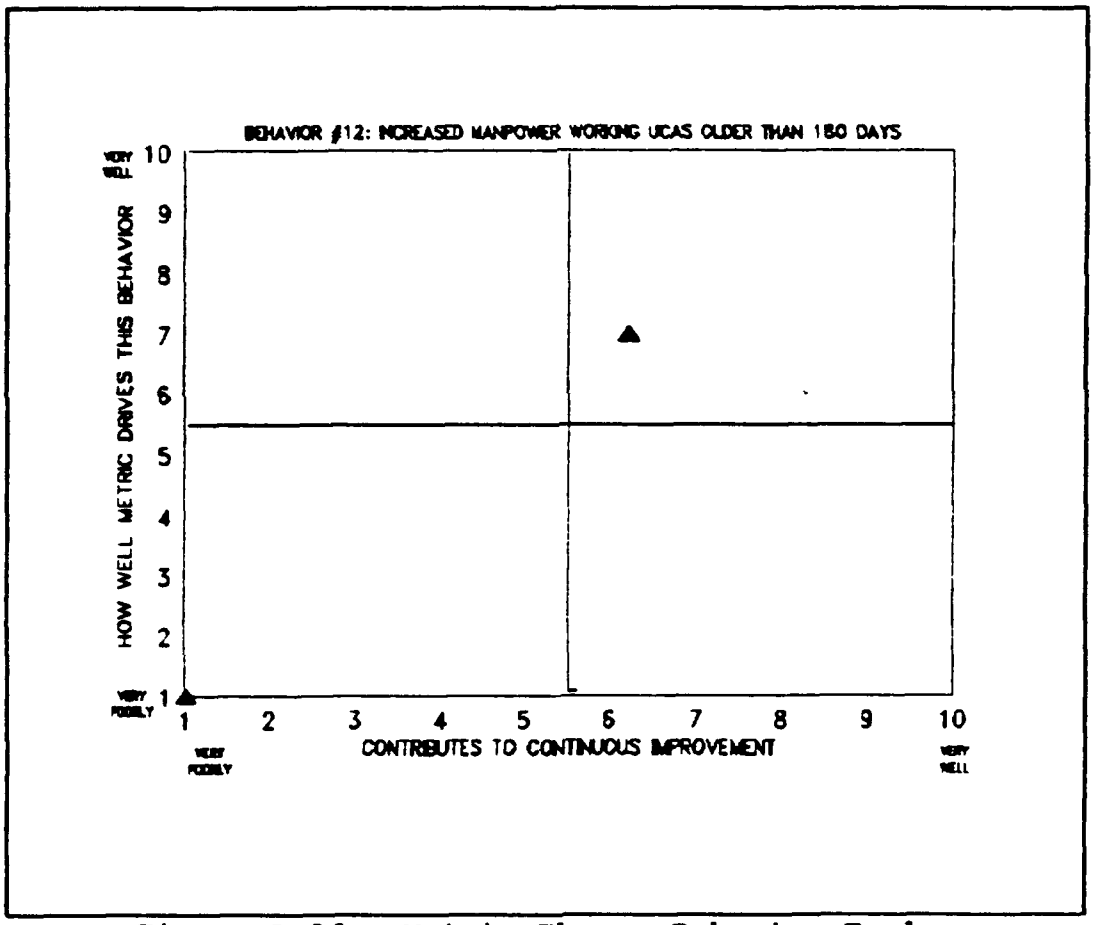


Figure A-36. Metric Three, Behavior Twelve

Behavior 13. Quicker negotiations

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	3	-	1	1	-	7.00
How well does behavior contribute to CI?	-	1	-	-	1	1	1	-	-	1	6.00

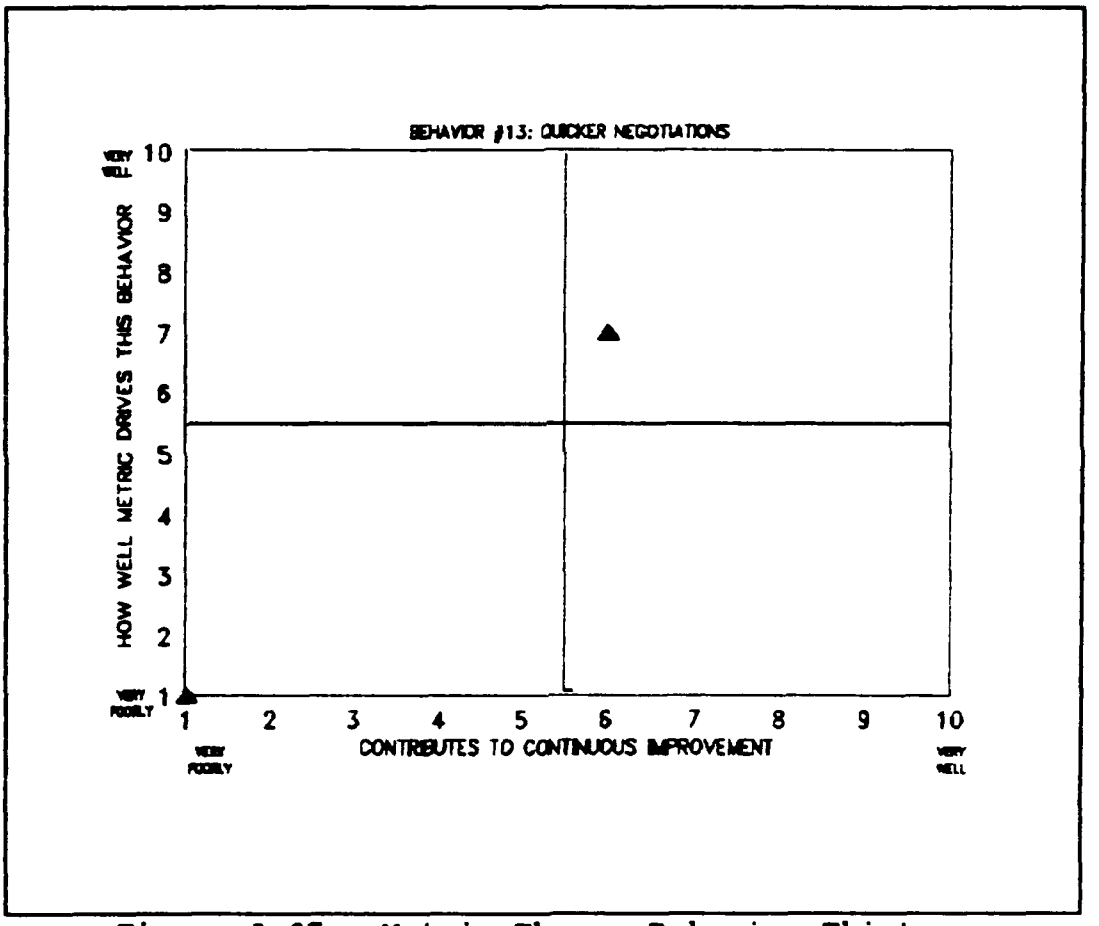


Figure A-37. Metric Three, Behavior Thirteen

Behavior 14. Other program activities placed on hold to finish definitization actions

**. Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	2	1	1	-	7.40
How well does behavior contribute to CI?	1	2	1	-	-	-	-	-	1	-	3.40

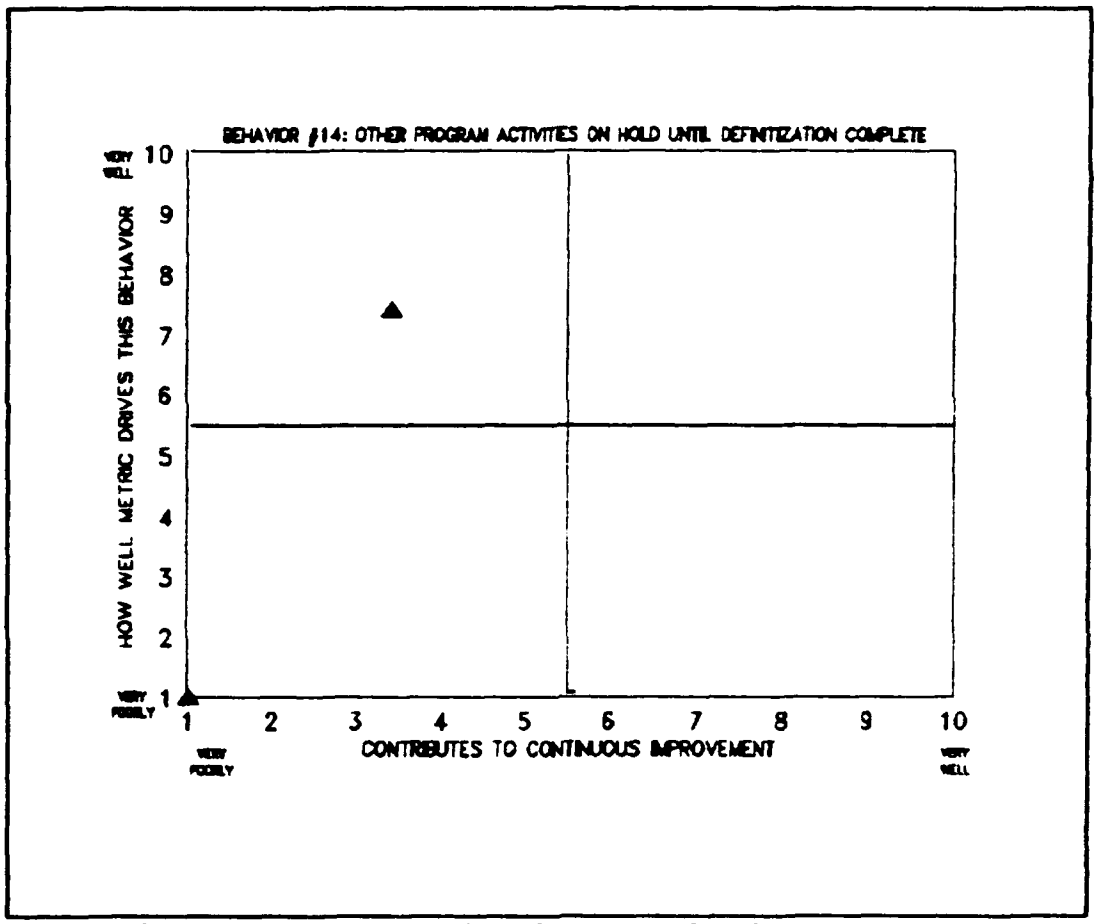


Figure A-38. Metric Three, Behavior Fourteen

Behavior 15. Contractor/SPO partners perform more up-front agreement on required actions

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	2	-	1	1	1	-	6.80
How well does behavior contribute to CI?	-	-	-	-	-	-	-	1	2	2	9.20

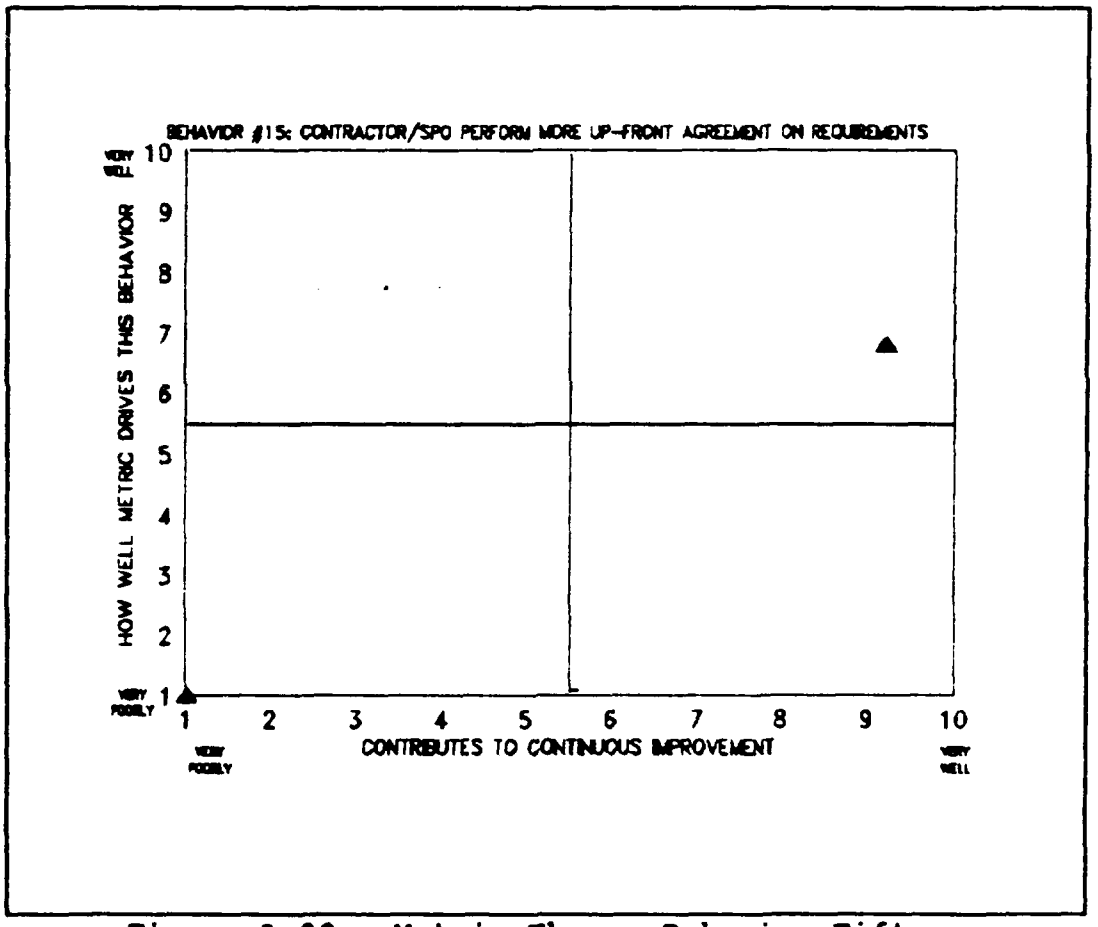


Figure A-39. Metric Three, Behavior Fifteen

Metric 4 Behavior 1. SPO will work to decrease number of breaches

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric	-	-	-	-	-	-	3	-	-	2	8.20
How well does behavior contribute to CI?	-	-	-	-	-	-	1	-	2	2	9.00

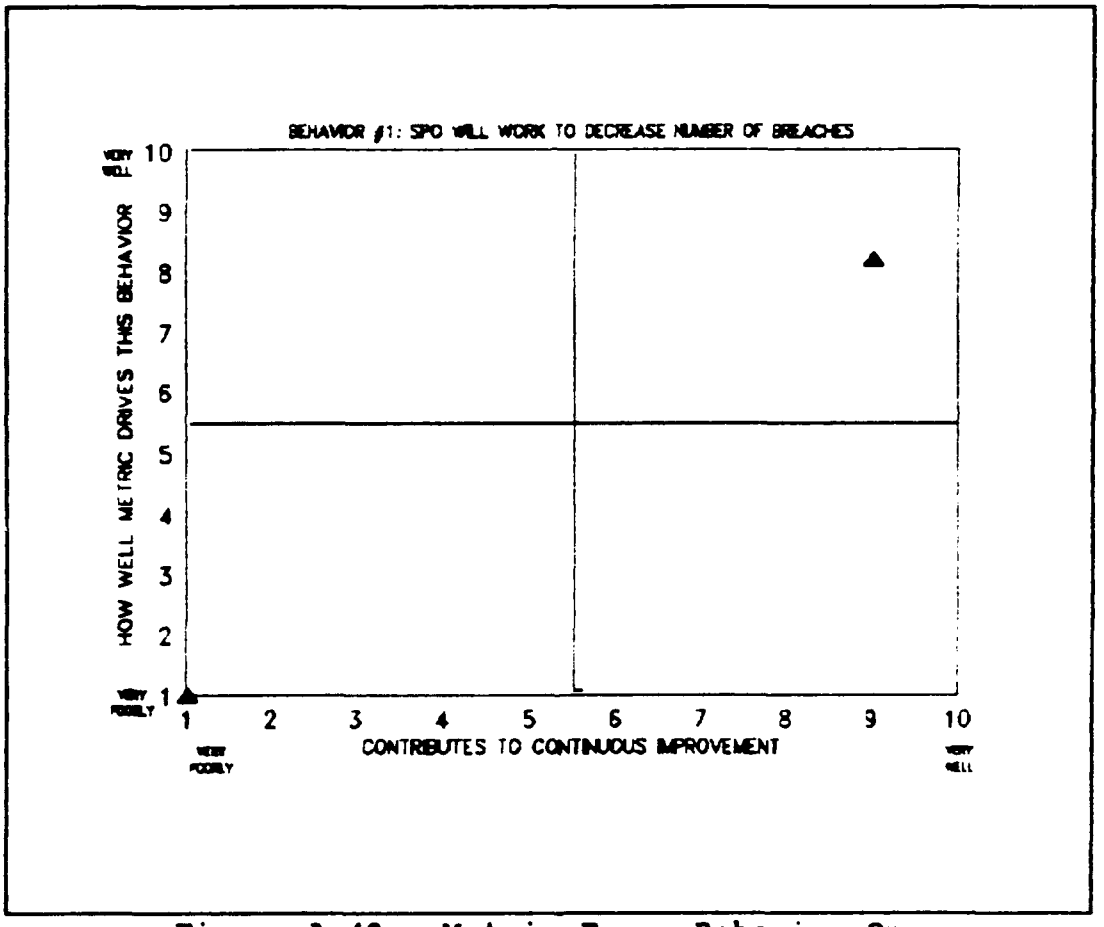


Figure A-40. Metric Four, Behavior One

Behavior 2. Project managers will interpret their own data as non-breaches

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	1	4	-	-	7.80
How well does behavior contribute to CI?	1	3	1	-	-	-	-	-	-	-	2.00

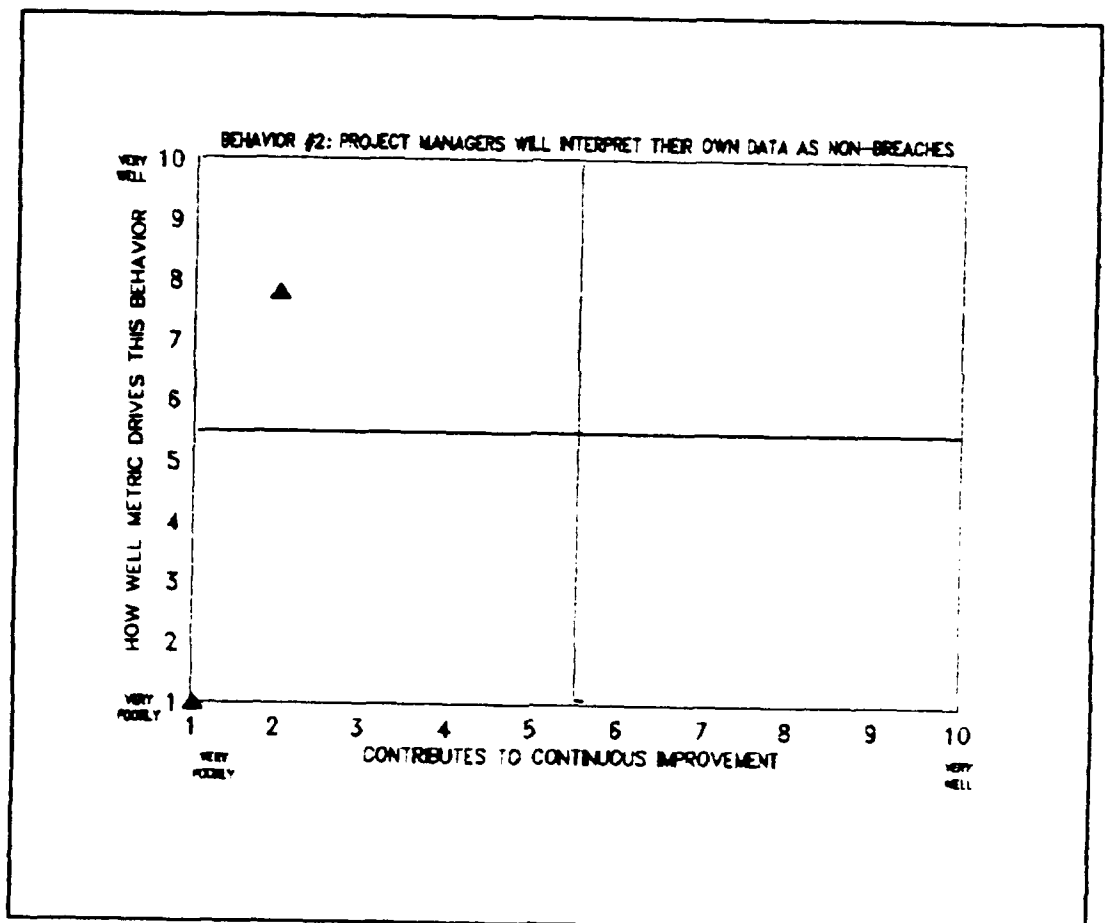


Figure A-41. Metric Four, Behavior Two

Behavior 3. SPO project team will set more conservative baselines

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	-	2	2	-	8.00
How well does behavior contribute to CI?	1	1	1	2	-	-	-	-	-	-	2.80

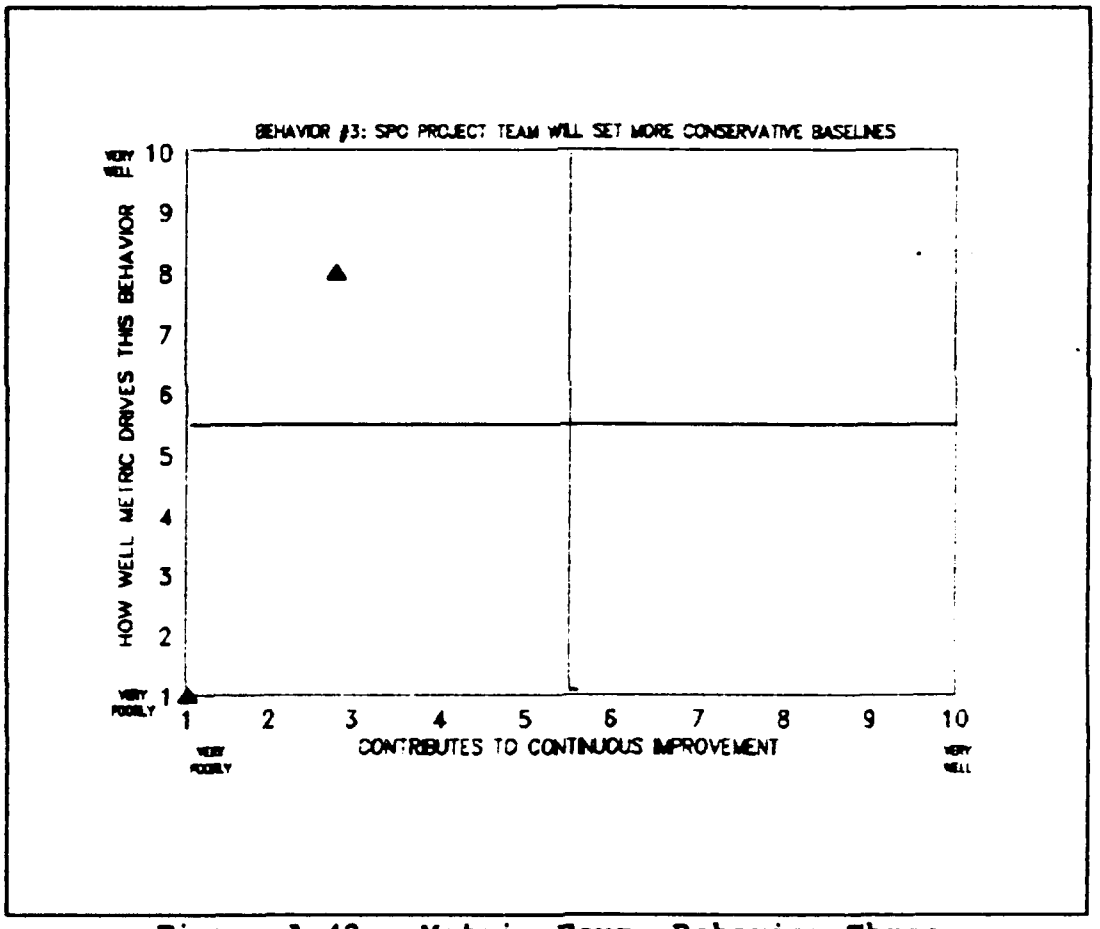


Figure A-42. Metric Four, Behavior Three

Behavior 4. Baselines written only on those actions w/high success possibilities

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	1	2	2	-	8.20
How well does behavior contribute to CI?	1	-	2	2	-	-	-	-	-	-	3.00

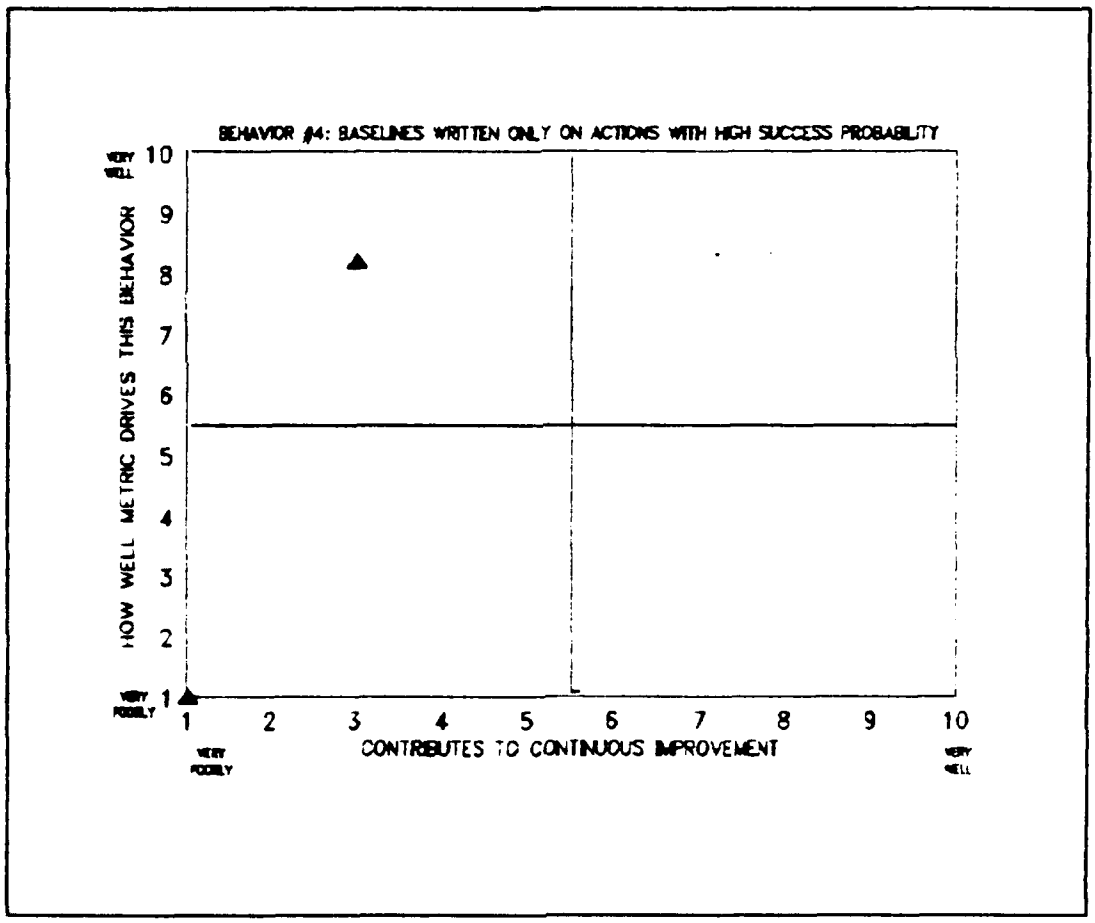


Figure A-43. Metric Four, Behavior Four

Behavior 5. Project team will write fewer baselines

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	-	4	-	-	7.60
How well does behavior contribute to CI?	1	-	2	1	-	1	-	-	-	-	3.40

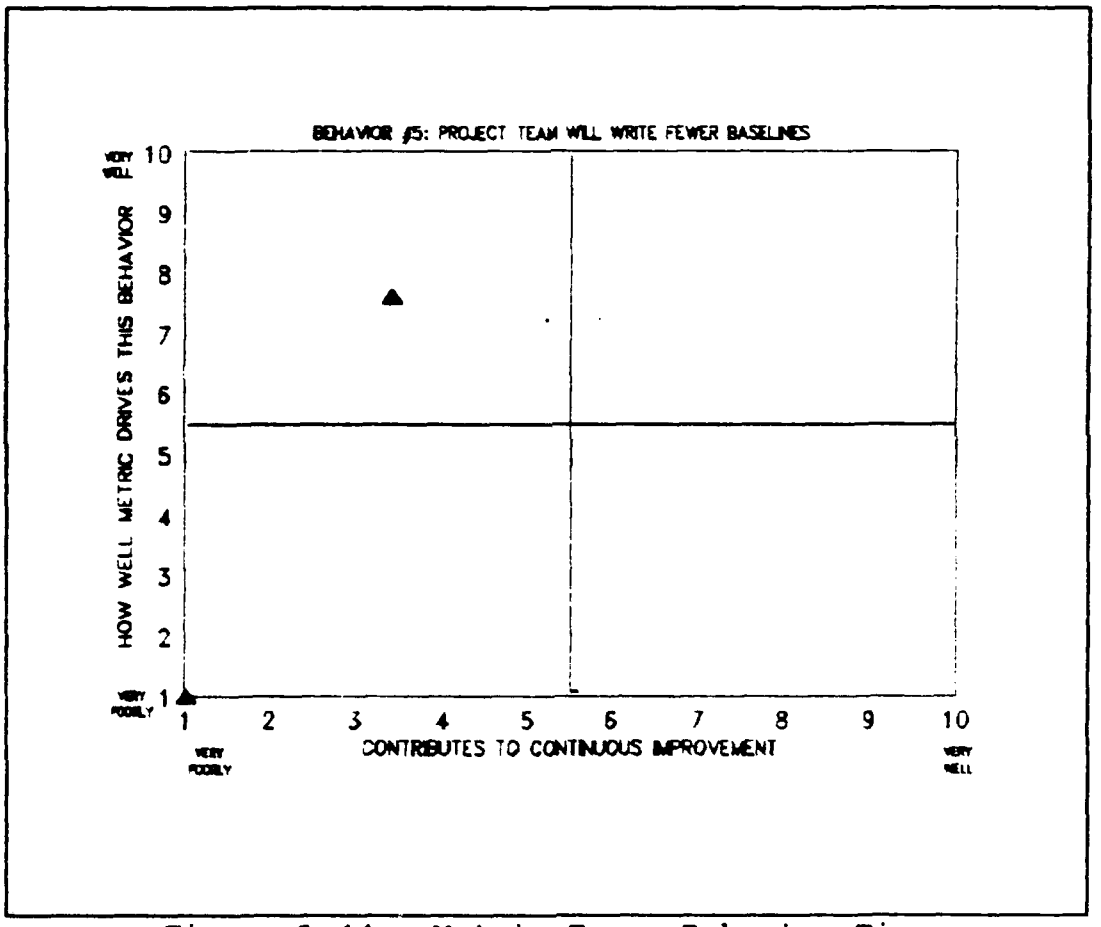


Figure A-44. Metric Four, Behavior Five

Behavior 6. Manpower will be shifted to work breached projects

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	2	1	1	-	7.40
How well does behavior contribute to CI?	-	1	-	-	-	1	1	1	1	-	6.40

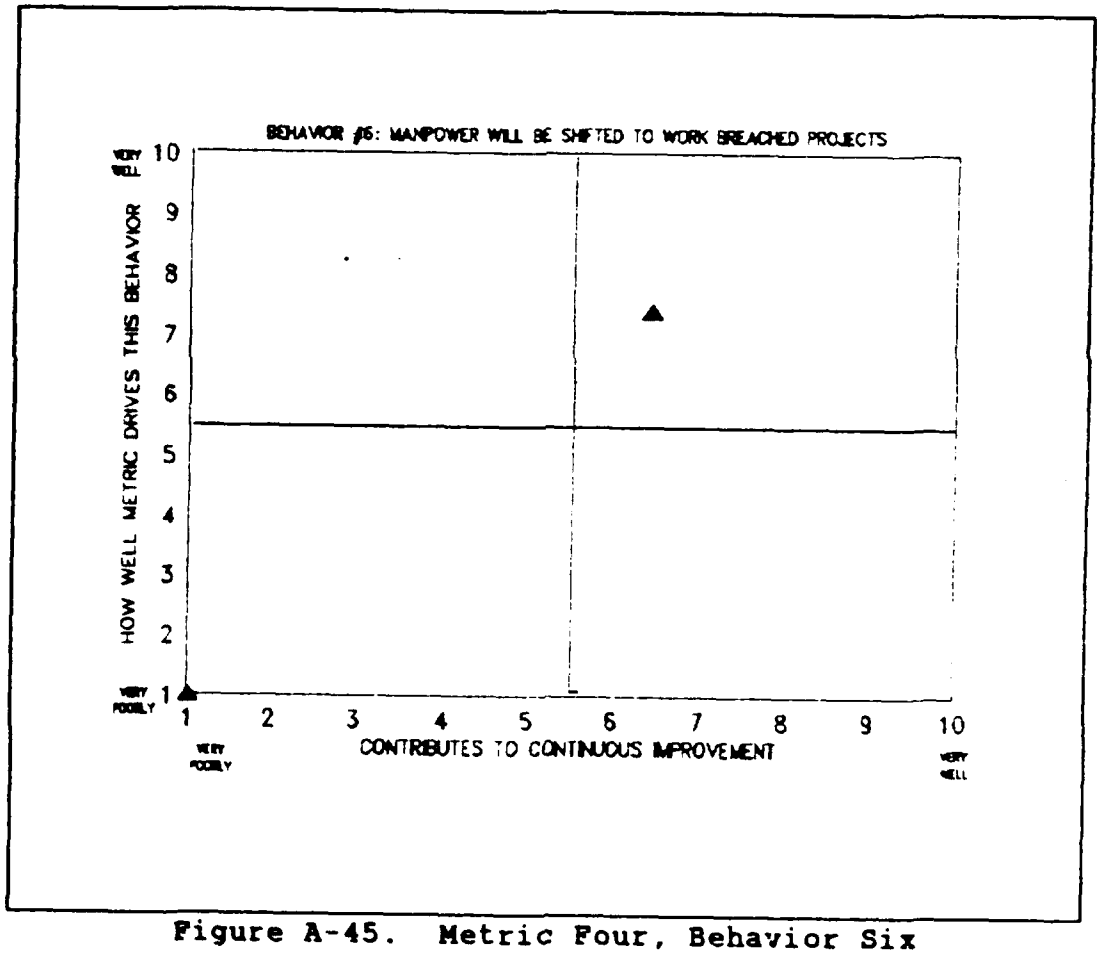


Figure A-45. Metric Four, Behavior Six

Behavior 7. Increased project status review/reporting

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	-	-	1	2	-	1	7.40
How well does behavior contribute to CI?	-	-	1	1	-	-	2	-	-	1	6.20

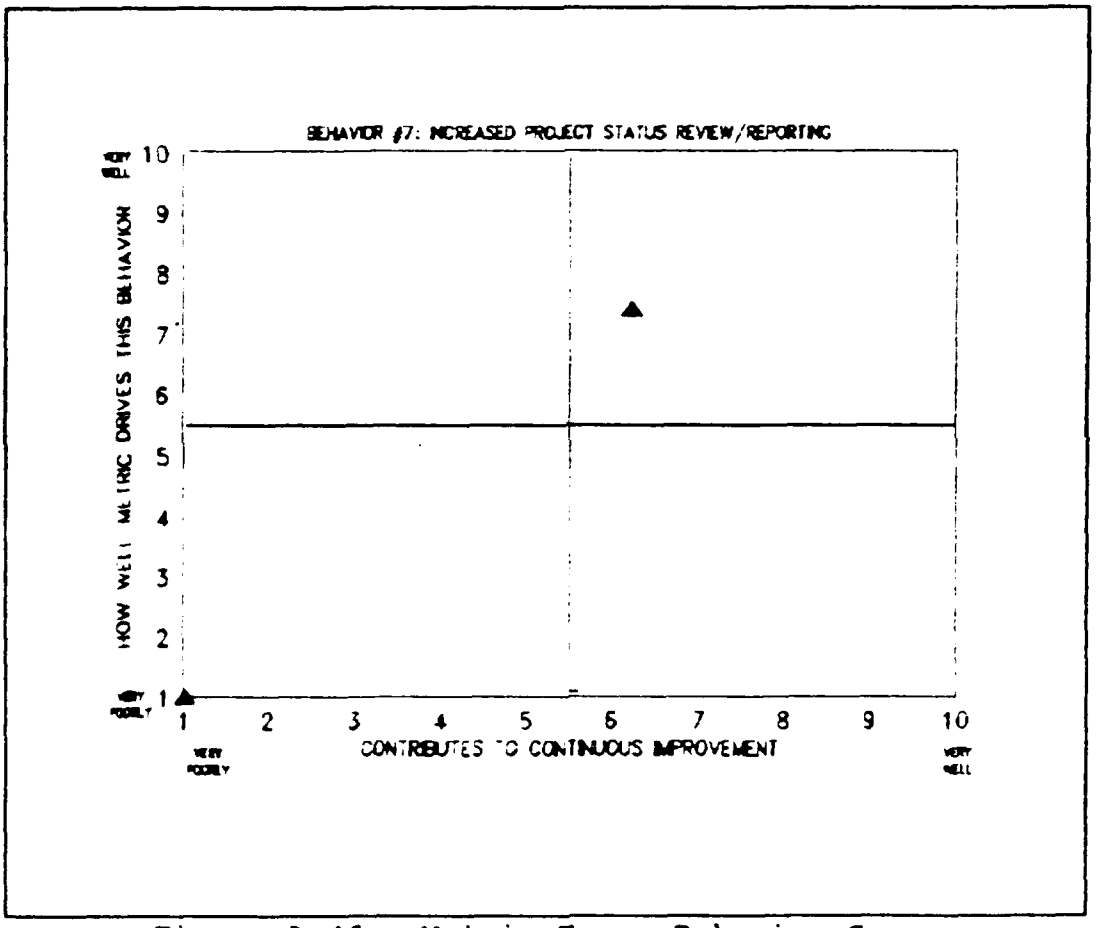


Figure A-46. Metric Four, Behavior Seven

Behavior 8. Upper management will help breached project managers

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	2	-	-	6.80
How well does behavior contribute to CI?	-	-	1	-	1	-	1	2	-	-	6.20

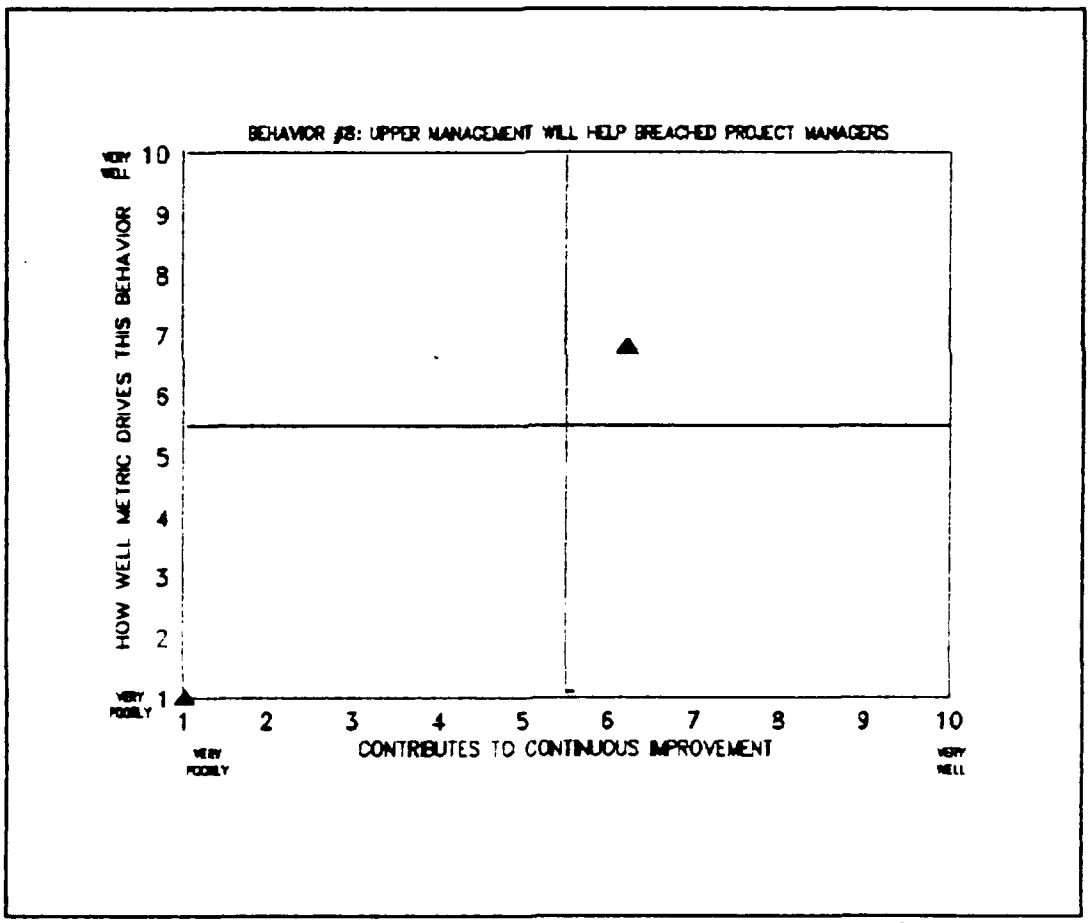


Figure A-47. Metric Four, Behavior Eight

Behavior 9. SPO will work individual breaches before entire program is breached

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	1	1	-	-	1	6.40
How well does behavior contribute to CI?	-	-	-	-	1	2	-	1	1	-	6.80

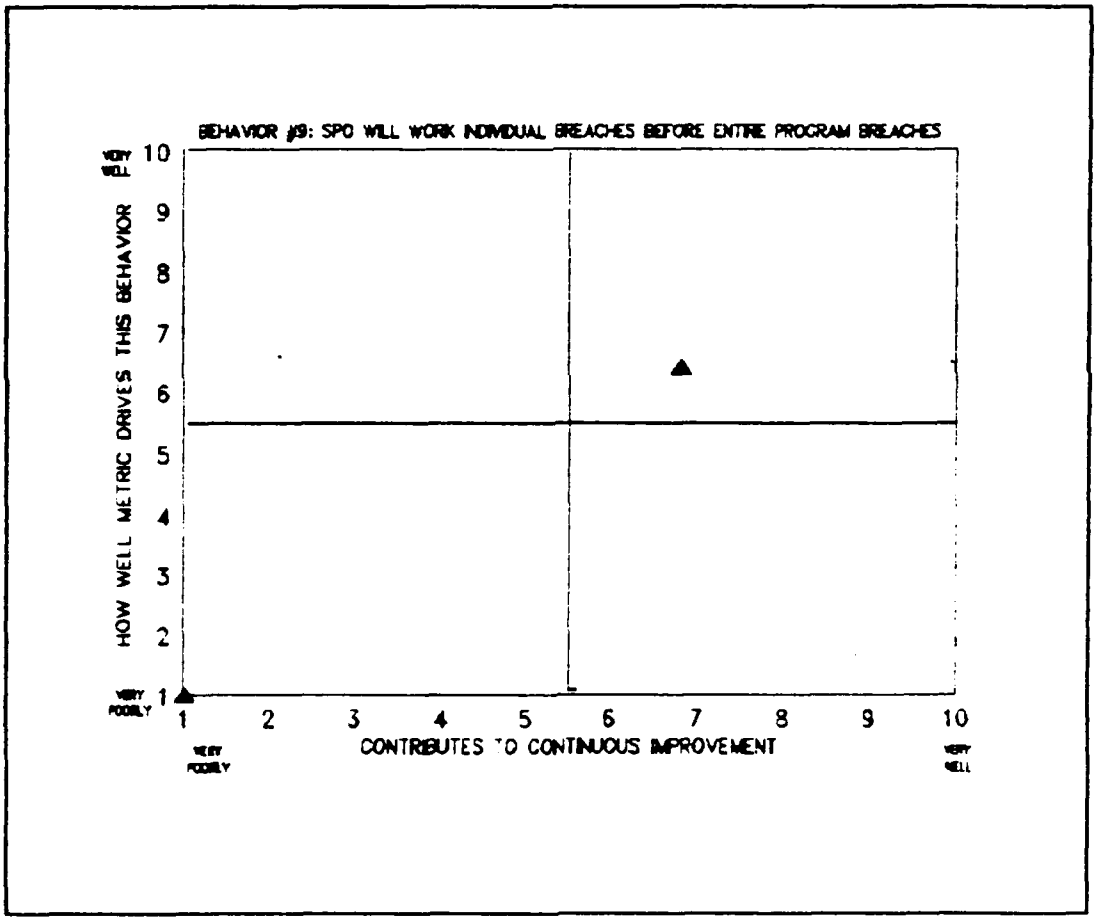


Figure A-48. Metric Four, Behavior Nine

Behavior 10. PM emphasize the term "informal" in order to explain away breaches.

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	1	-	-	-	-	1	2	1	-	6.80
How well does behavior contribute to CI?	1	3	1	-	-	-	-	-	-	-	2.00

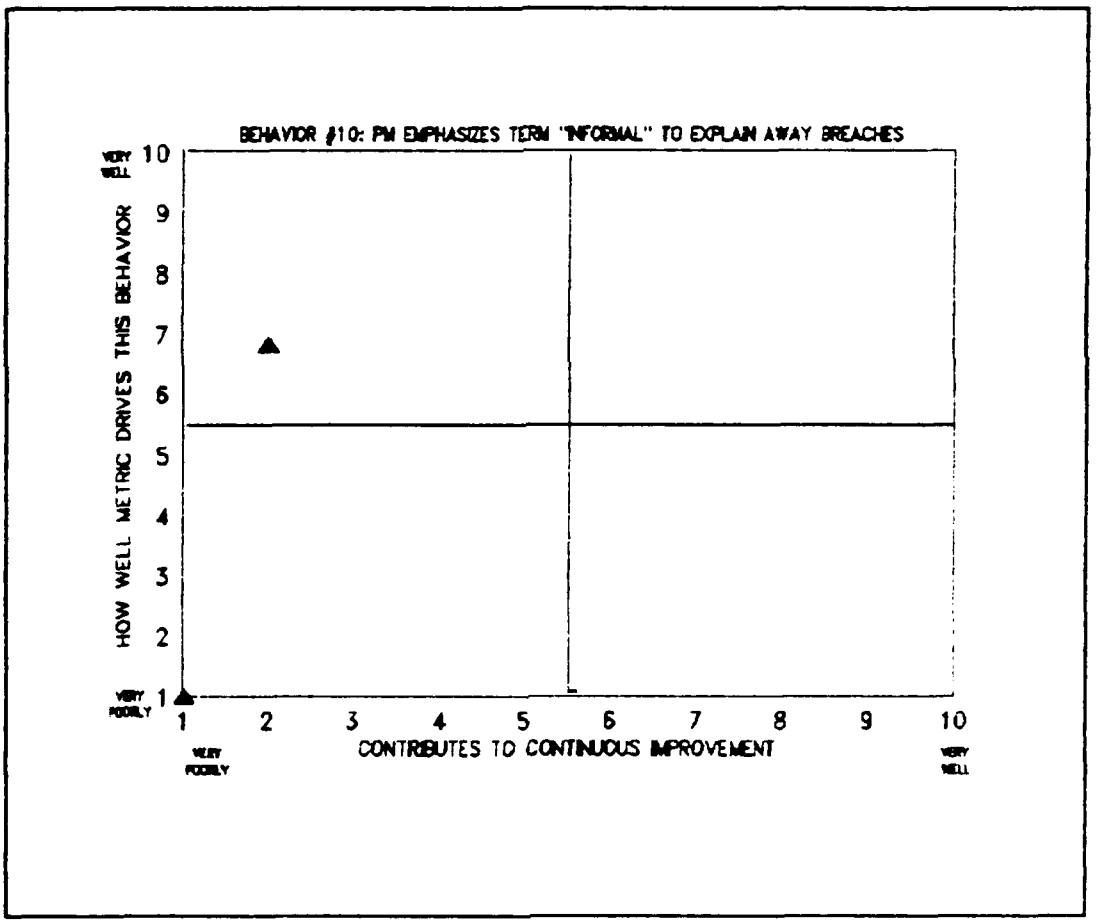


Figure A-49. Metric Four, Behavior Ten

Behavior 11. Monthly reporting will shift management emphasis to short-term problems

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	2	3	-	-	7.60
How well does behavior contribute to CI?	-	2	1	1	-	1	-	-	-	-	3.40

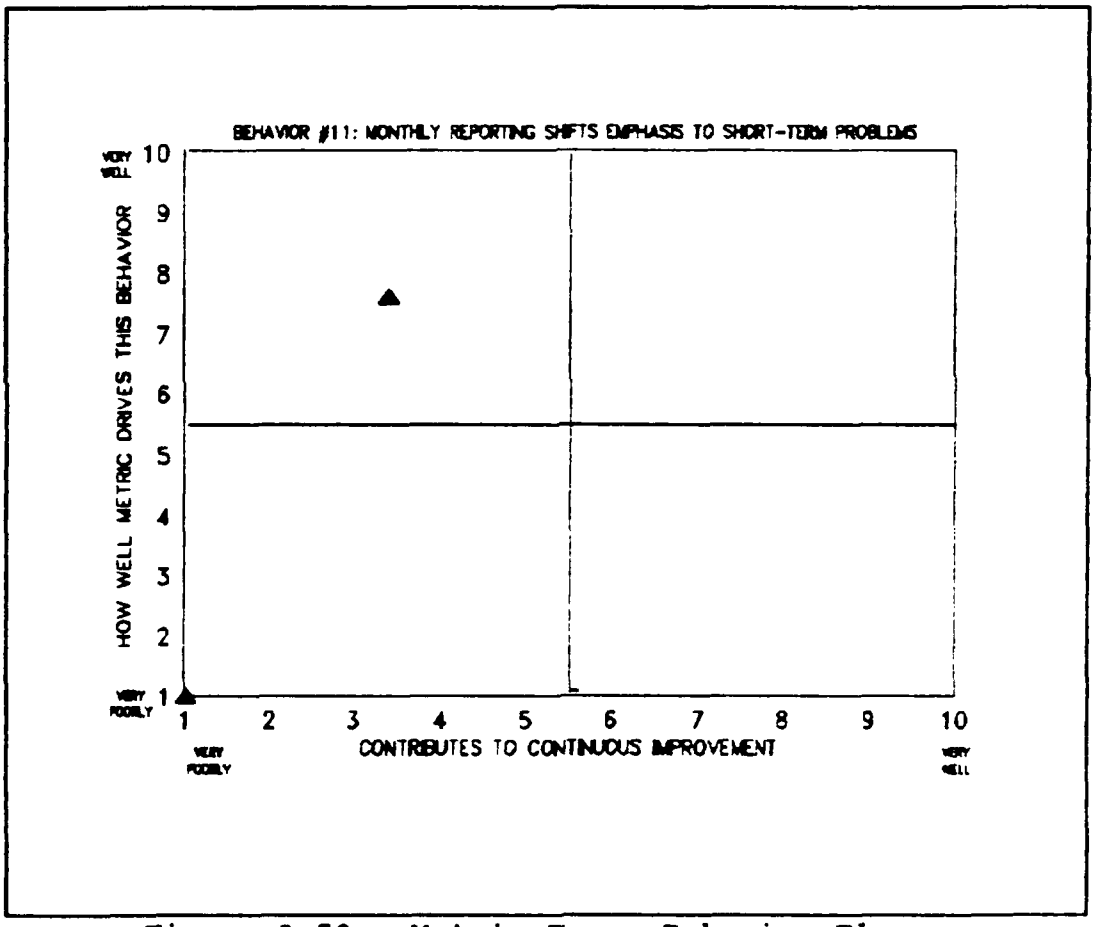


Figure A-50. Metric Four, Behavior Eleven

Behavior 12. Finger pointing increases

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	2	-	1	-	-	2	-	-	5.40
How well does behavior contribute to CI?	2	2	1	-	-	-	-	-	-	-	1.80

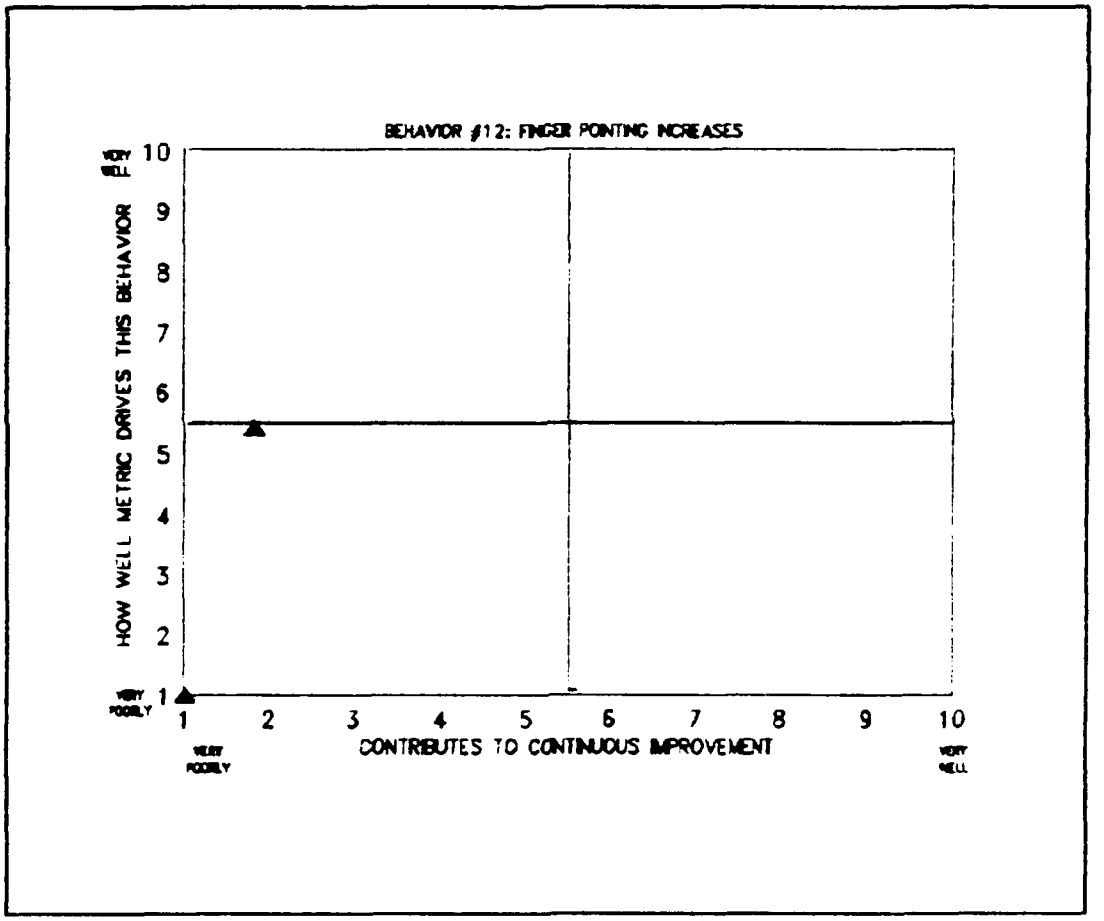


Figure A-51. Metric Four, Behavior Twelve

Behavior 13. Problems will be worked according to probability of causing breach

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	-	3	-	-	7.00
How well does behavior contribute to CI?	-	-	1	1	2	-	-	1	-	-	5.00

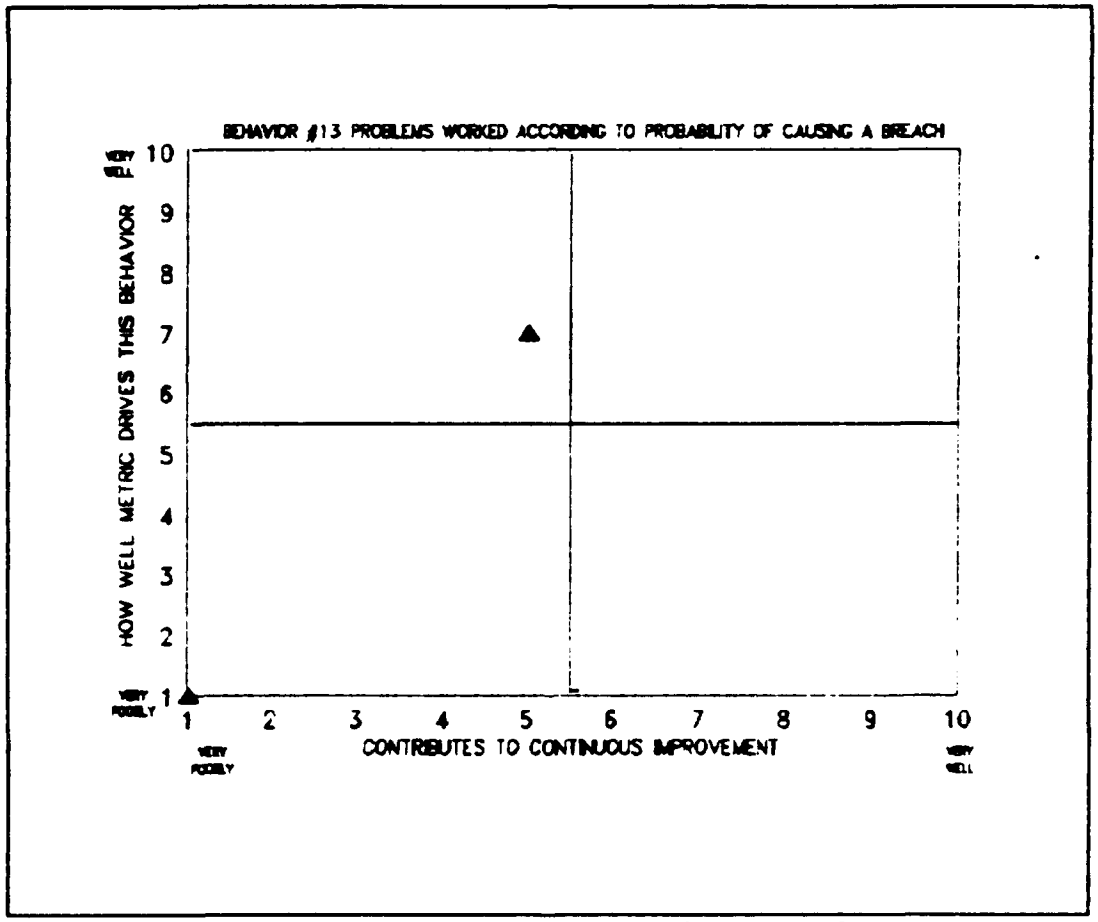


Figure A-52. Metric Four, Behavior Thirteen

Behavior 14. Easy-to-work breaches will be solved before tough ones

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	2	1	1	1	-	-	6.20
How well does behavior contribute to CI?	-	1	-	1	2	1	-	-	-	-	4.40

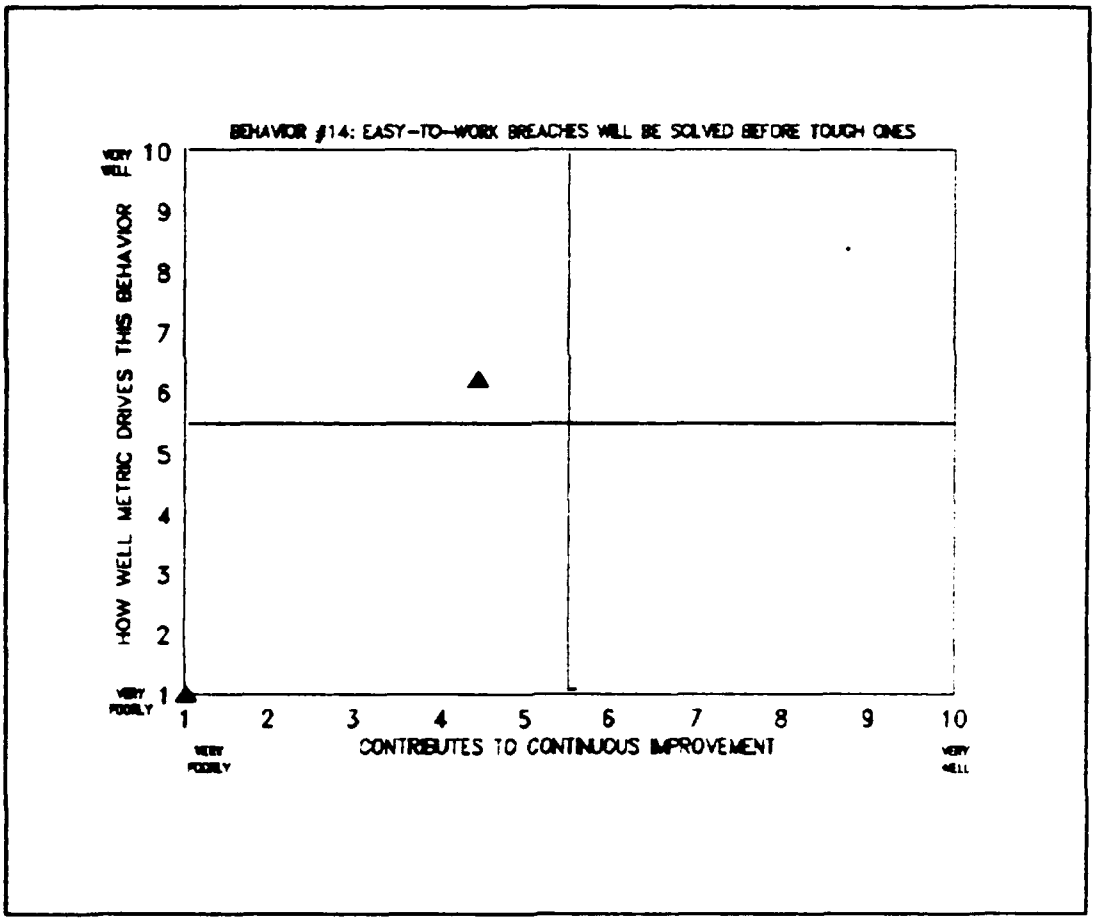


Figure A-53. Metric Four, Behavior Fourteen

Behavior 15. PMs will divide work according to number of potential breaches

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	2	-	2	-	-	-	5.60
How well does behavior contribute to CI?	-	-	2	2	1	-	-	-	-	-	3.80

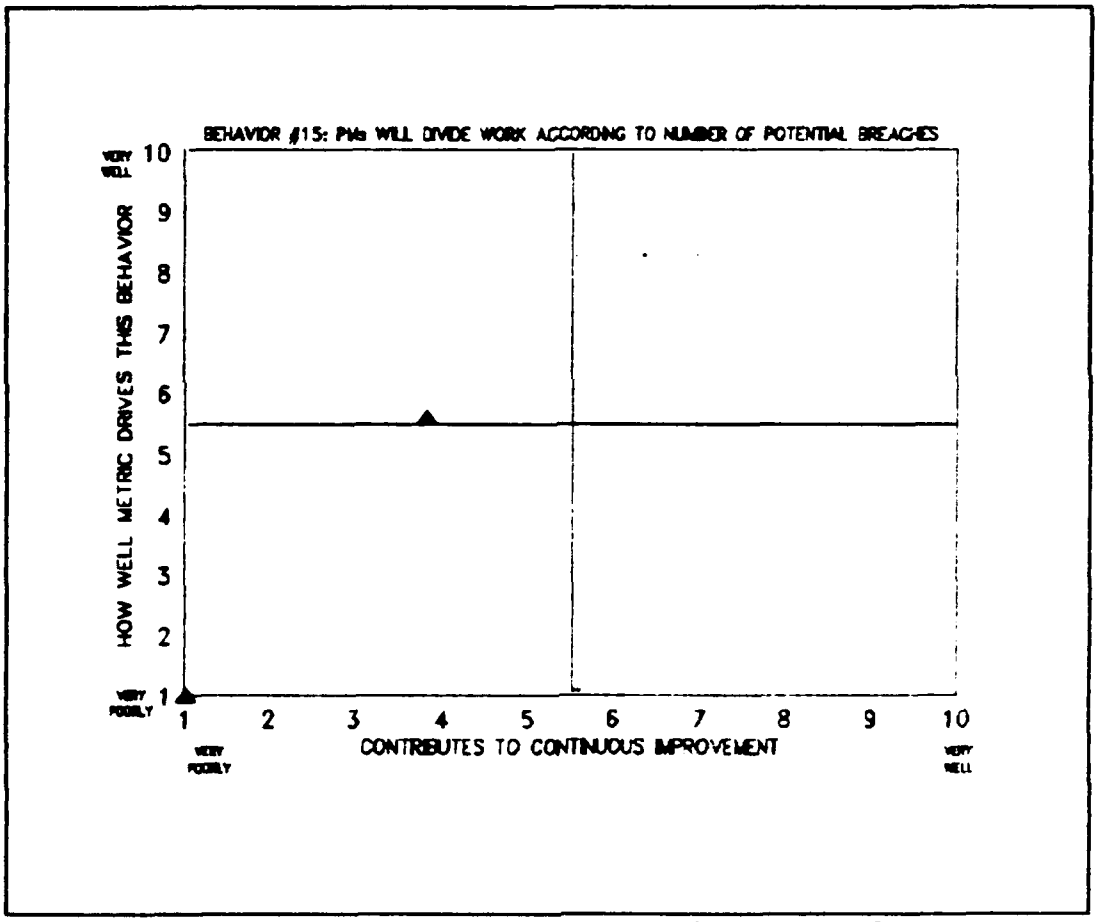


Figure A-54. Metric Four, Behavior Fifteen

Behavior 16. PM focuses on number of breaches rather than reasons/severity for breaches

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behaivor?	-	-	-	1	-	1	1	1	1	-	6.80
How well does behavior contribute to CI?	1	1	2	-	-	-	-	1	-	-	3.40

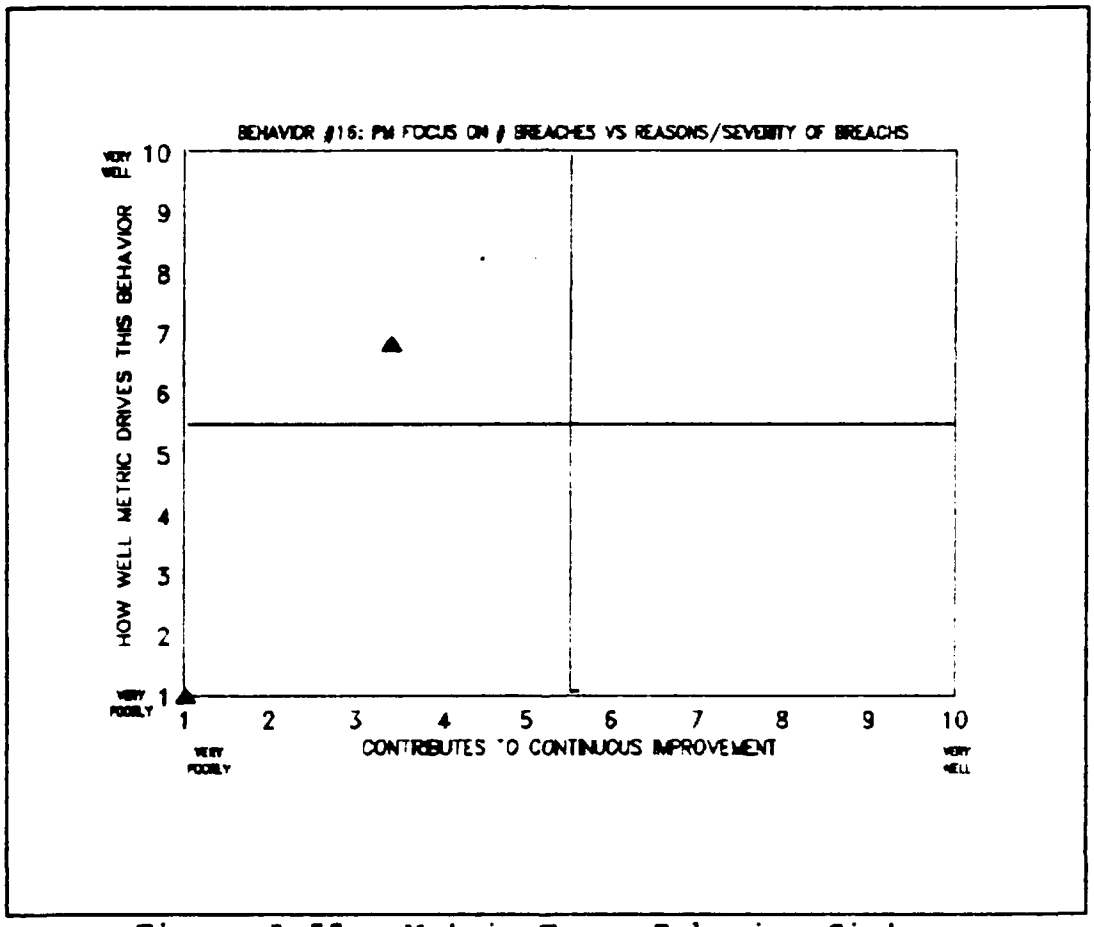


Figure A-55. Metric Four, Behavior Sixteen

Metric 5 Behavior 1. SPO will make conservative estimates of scheduled release

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	1	1	1	2	8.80
How well does behavior contribute to CI?	-	1	4	-	-	-	-	-	-	-	2.80

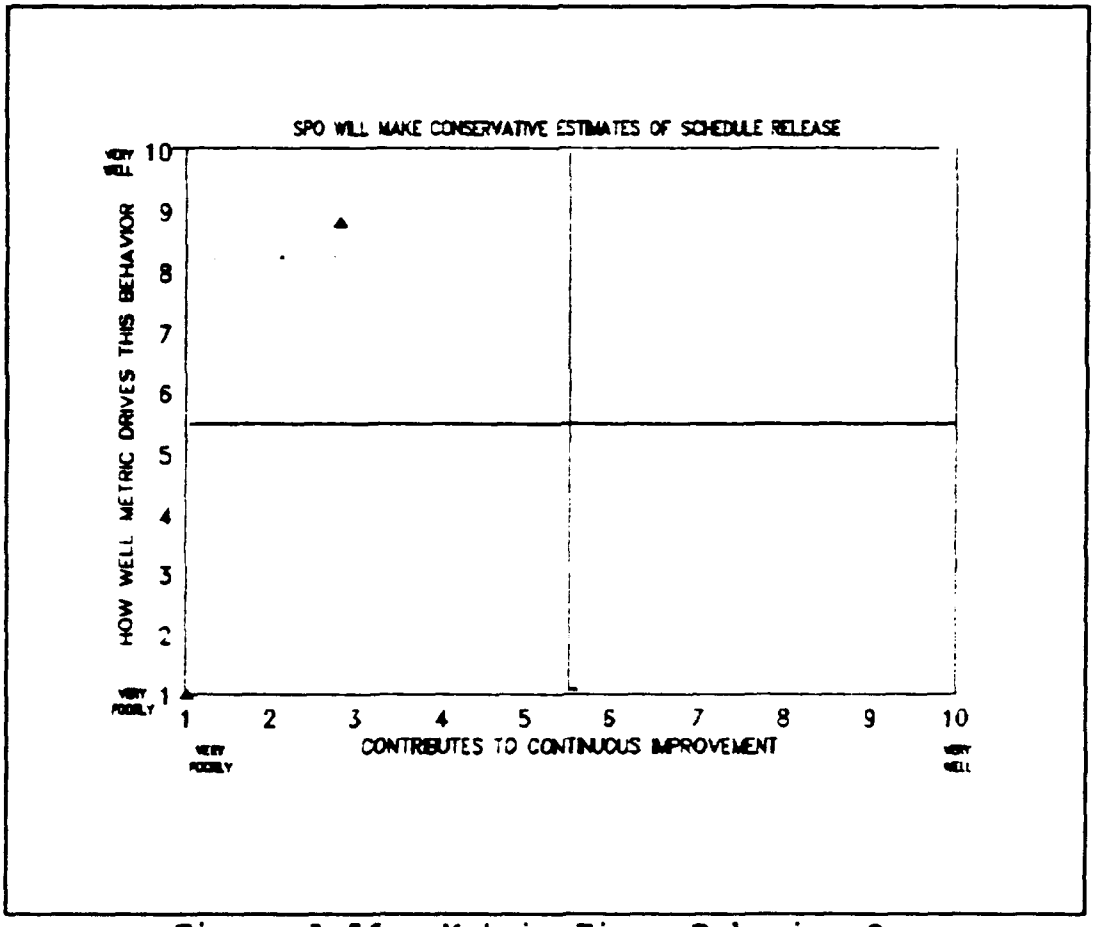


Figure A-56. Metric Five, Behavior One

Behavior 2. SPO will concentrate on making schedule rather than a good RFP

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	2	3	-	-	7.60
How well does behavior contribute to CI?	1	2	2	-	-	-	-	-	-	-	2.20

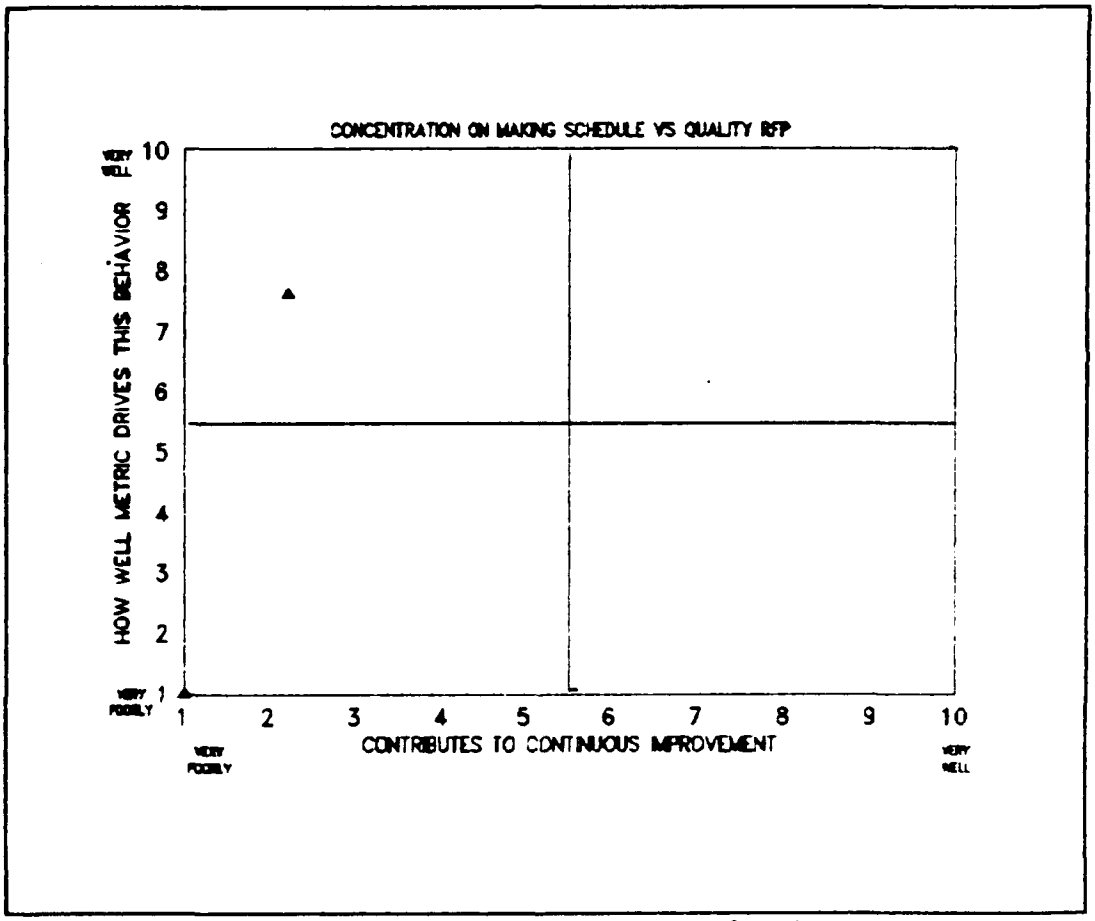


Figure A-57. Metric Five, Behavior Two

Behavior 3. Determine the components of the RFP cycle

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	-	1	1	7.40
How well does behavior contribute to CI?	-	-	-	-	-	1	-	-	2	2	8.80

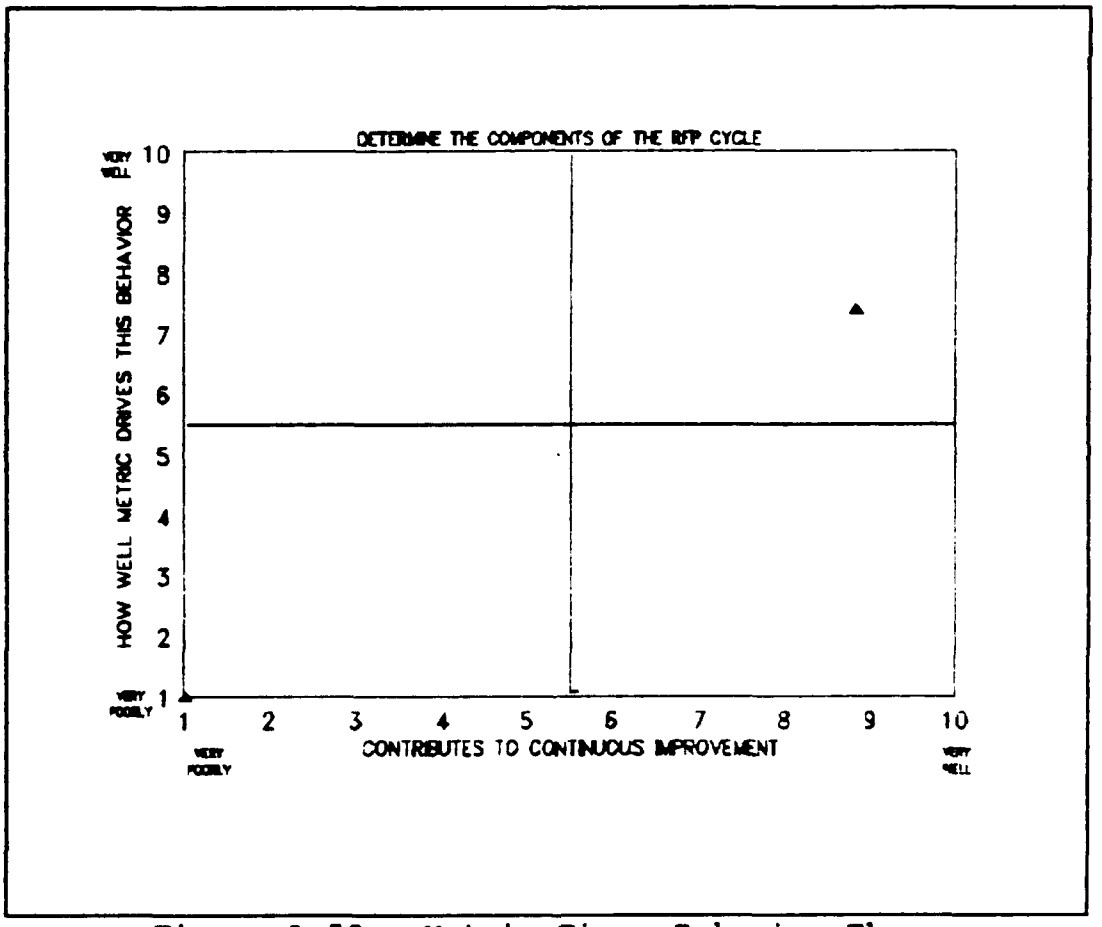


Figure A-58. Metric Five, Behavior Three

Behavior 4. SPO increases reviews on proposal development status

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	-	-	2	1	-	1	7.00
How well does behavior contribute to CI?	1	-	-	2	-	-	-	1	-	1	5.4

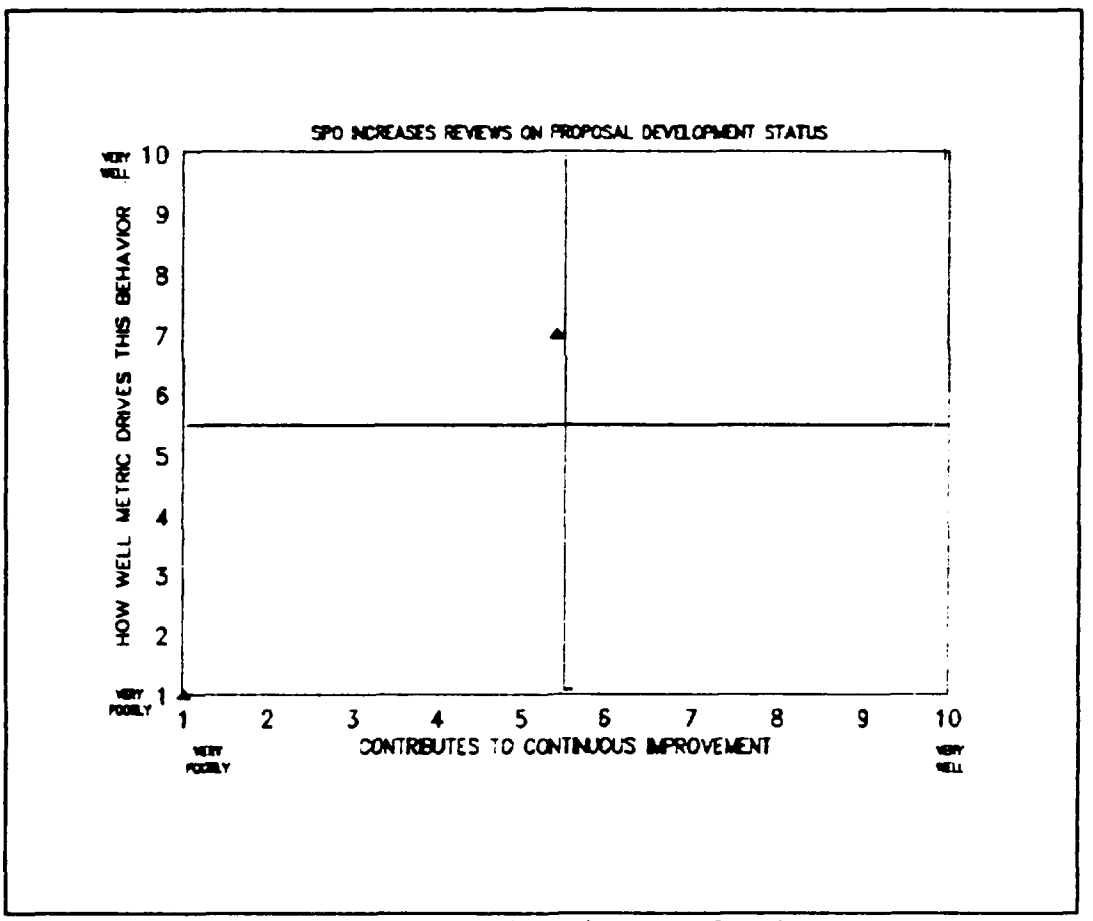


Figure A-59. Metric Five, Behavior Four

Behavior 5. SPO will develop a good process for RFP development

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	1	-	2	-	-	6.20
How well does behavior contribute to CI?	-	-	-	-	-	-	-	1	1	3	9.40

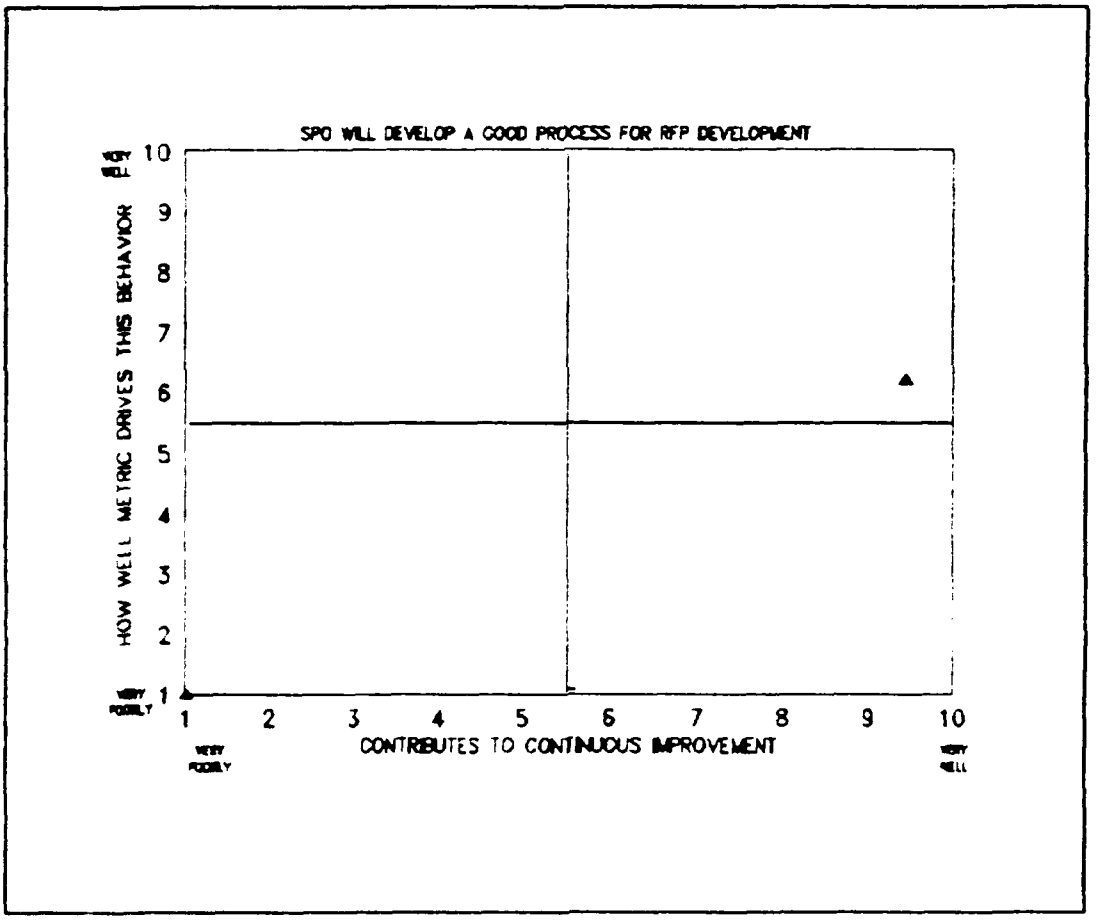


Figure A-60. Metric Five, Behavior Five

Behavior 6. Increased visibility on schedule intensive proposal development tasks

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	1	-	-	1	-	2	7.40
How well does behavior contribute to CI?	-	-	-	-	-	-	2	2	-	1	8.00

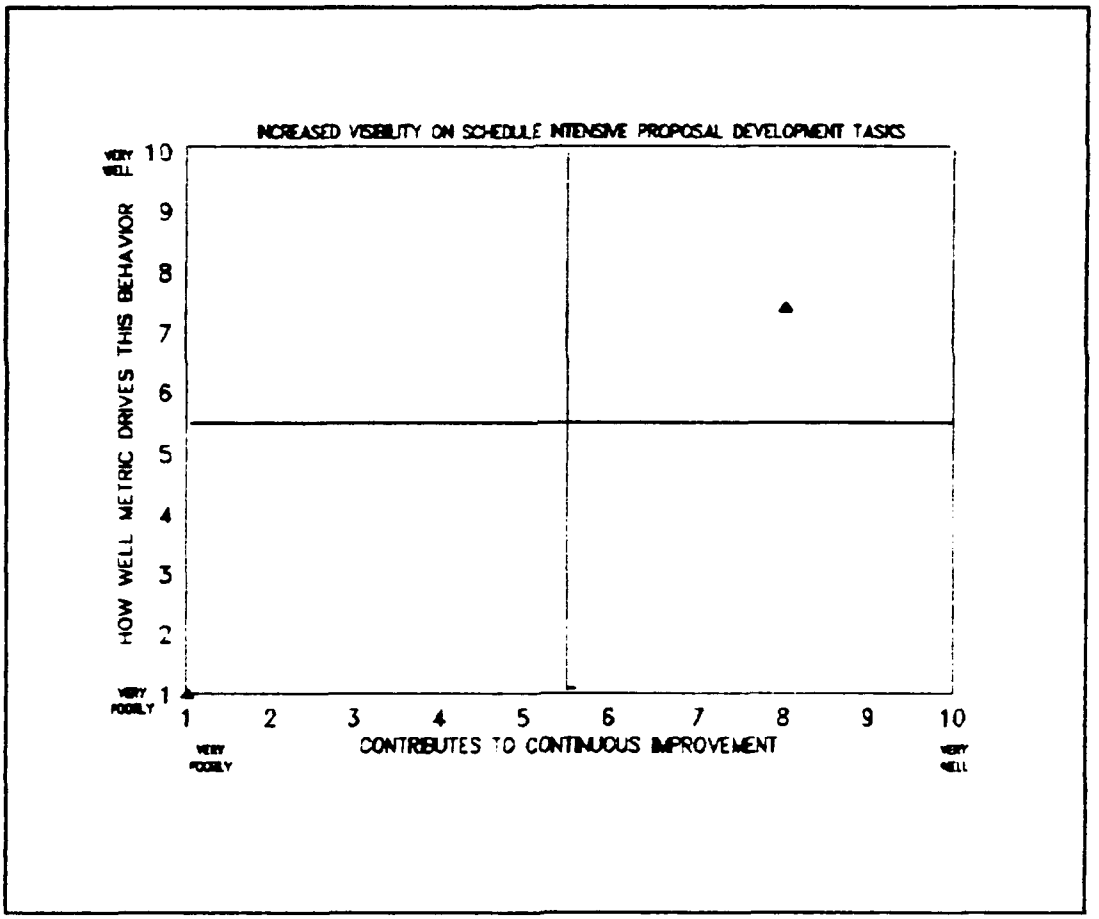


Figure A-61. Metric Five, Behavior Six

Behavior 7. SPO will pressure coordinating offices to shorten reviews

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	2	-	1	1	7.80
How well does behavior contribute to CI?	-	1	2	1	1	-	-	-	-	-	3.40

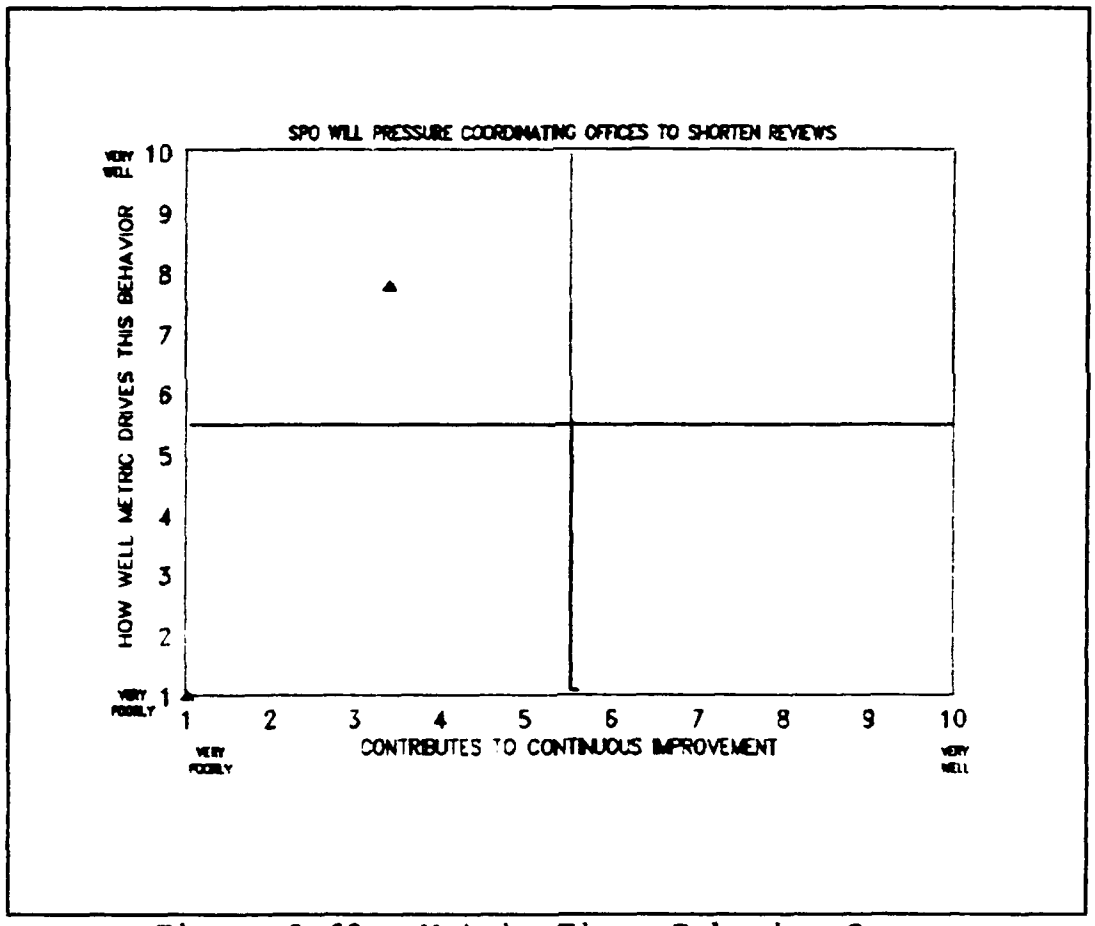


Figure A-62. Metric Five, Behavior Seven

Behavior 8. Manhours will be moved from other factions of program to work on RFPs/ECPs

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	2	-	-	6.80
How well does behavior contribute to CI?	-	-	2	1	1	-	1	-	-	-	4.40

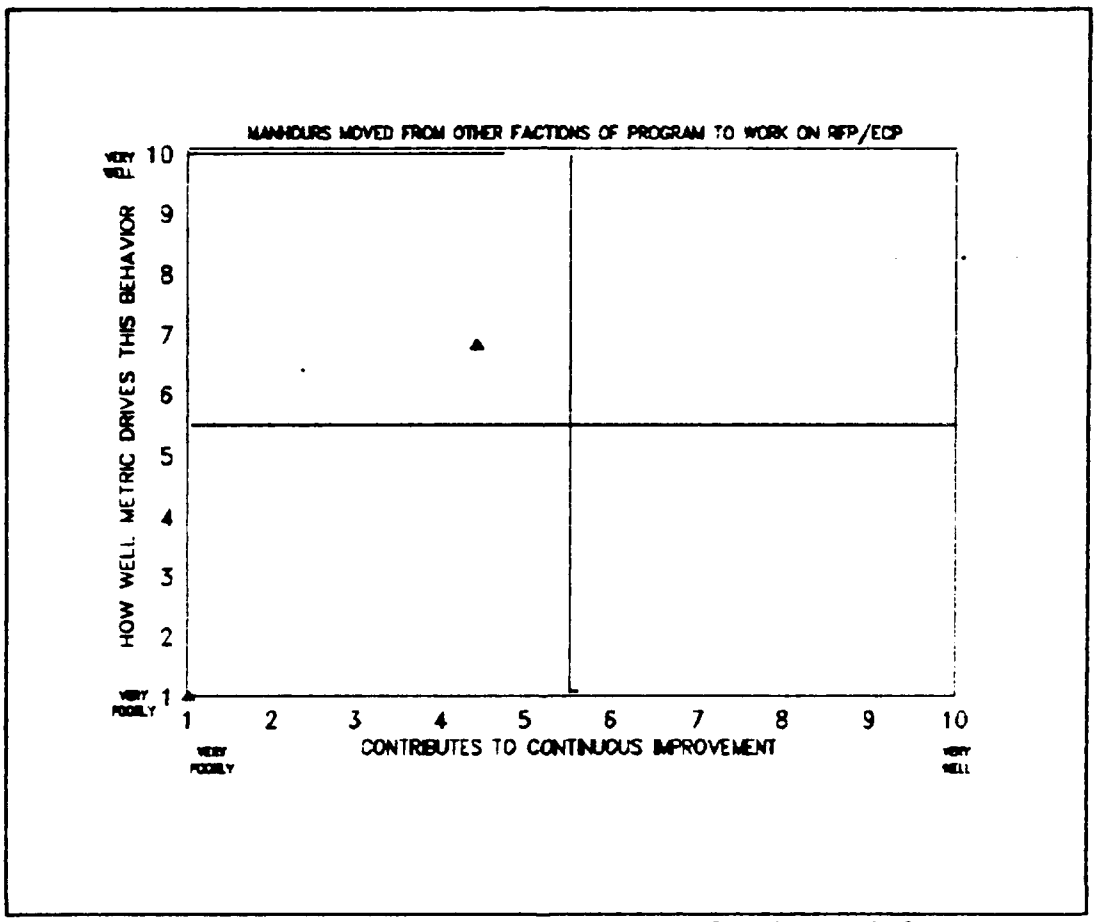


Figure A-63. Metric Five, Behavior Eight

Behavior 9. SPO will tend to use "canned" RFP packages

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	1	-	-	-	-	2	2	-	-	6.40
How well does behavior contribute to CI?	2	1	-	-	2	-	-	-	-	-	2.80

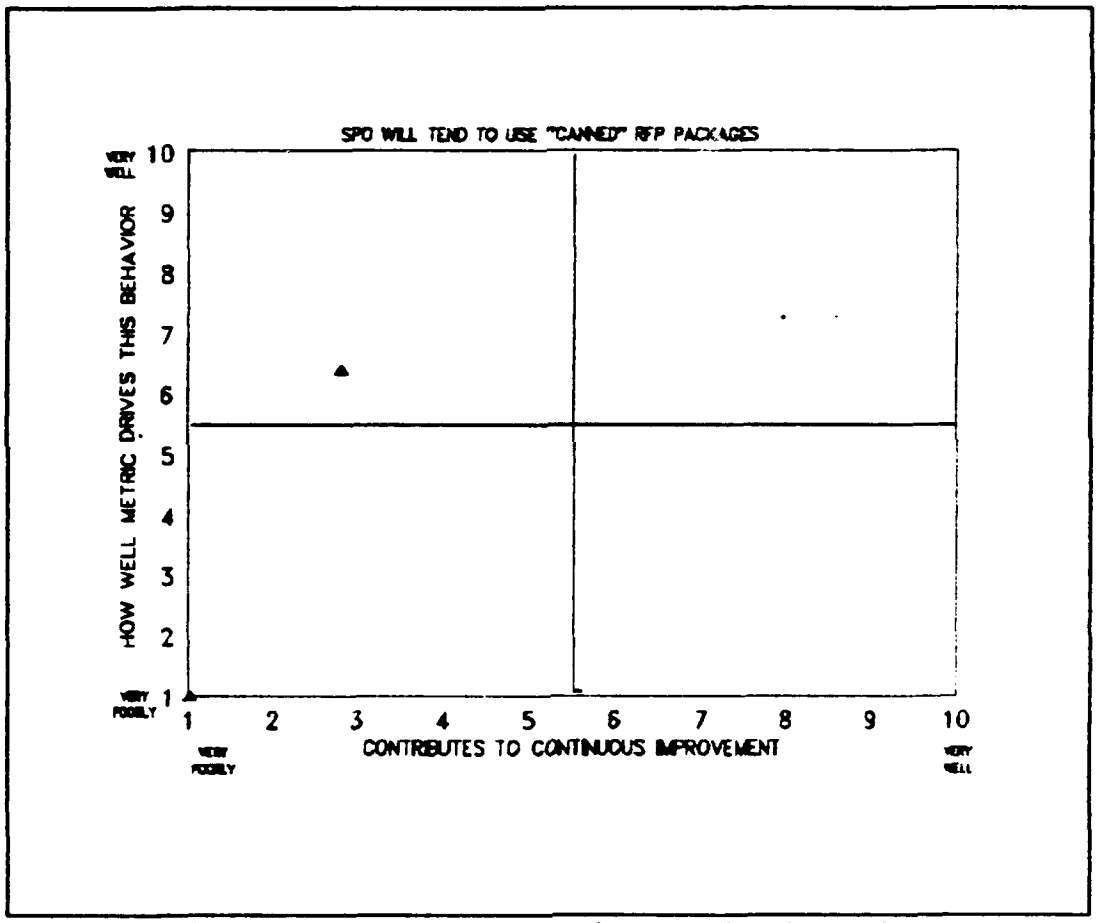


Figure A-64. Metric Five, Behavior Nine

Behavior 10. SPO/Contractor will do joint RFP development

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	3	1	1	-	-	-	-	4.60
How well does behavior contribute to CI?	-	-	-	-	-	1	1	2	1	-	7.60

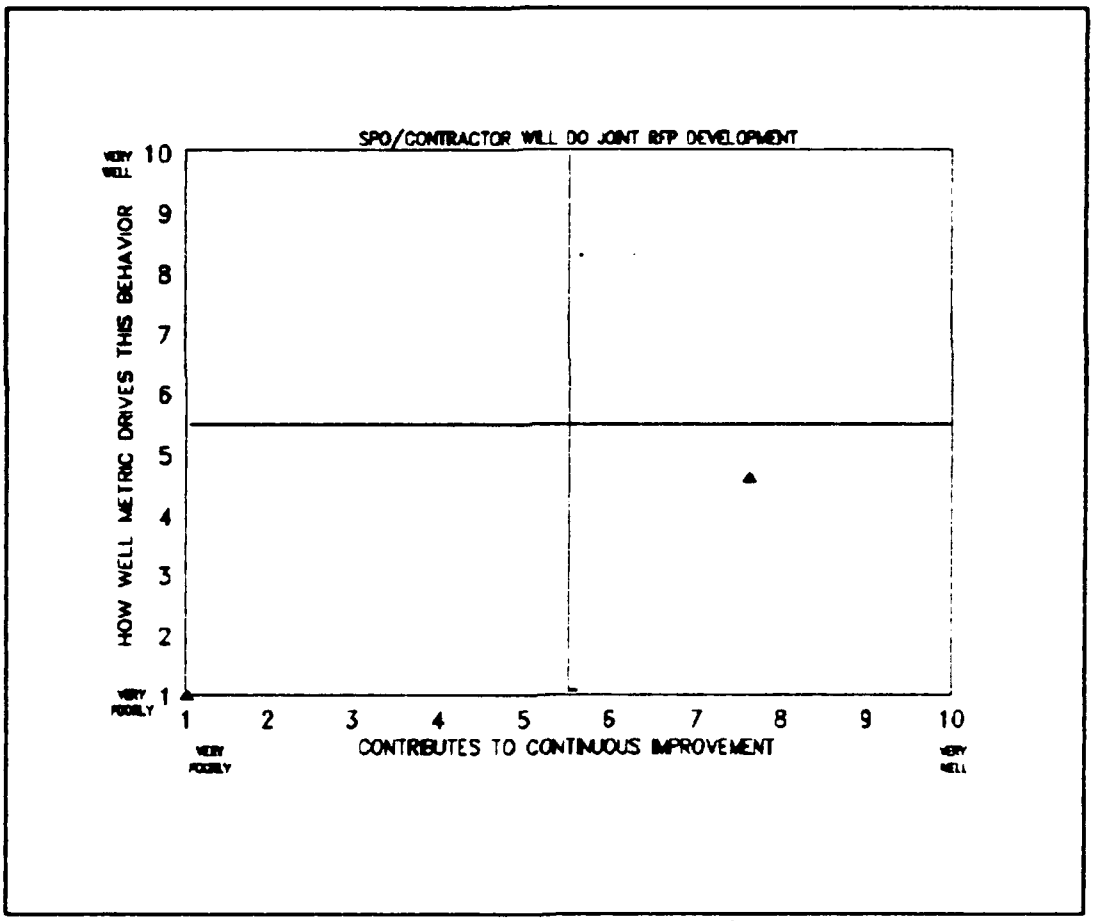


Figure A-65. Metric Five, Behavior Ten

Behavior 11. SPO will try to reduce the number of RFPs

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	4	1	-	-	7.20
How well does behavior contribute to CI?	1	-	1	1	2	-	-	-	-	-	3.60

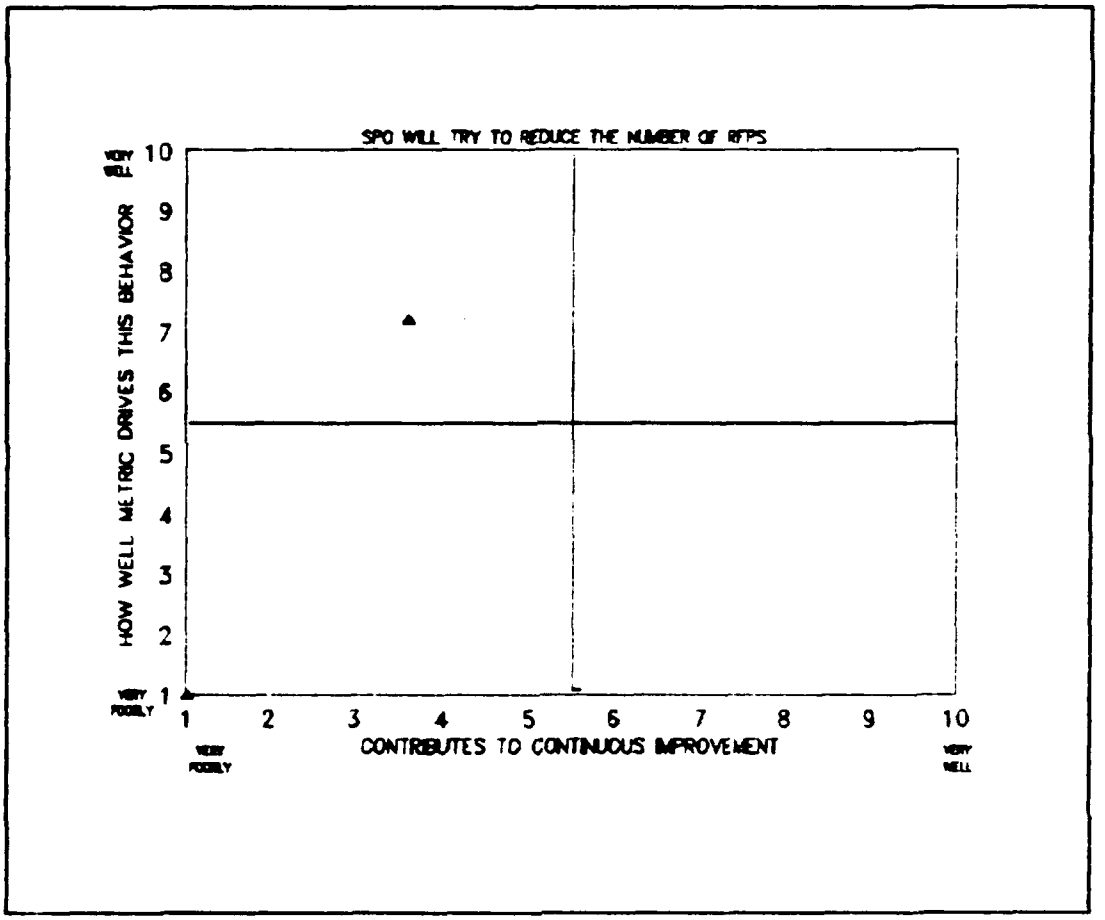


Figure A-66. Metric Five, Behavior Eleven

Behavior 12. RFP prep priorities will be based upon the ability to meet schedules

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	1	2	1	-	7.60
How well does behavior contribute to CI?	-	-	3	1	1	-	-	-	-	-	3.60

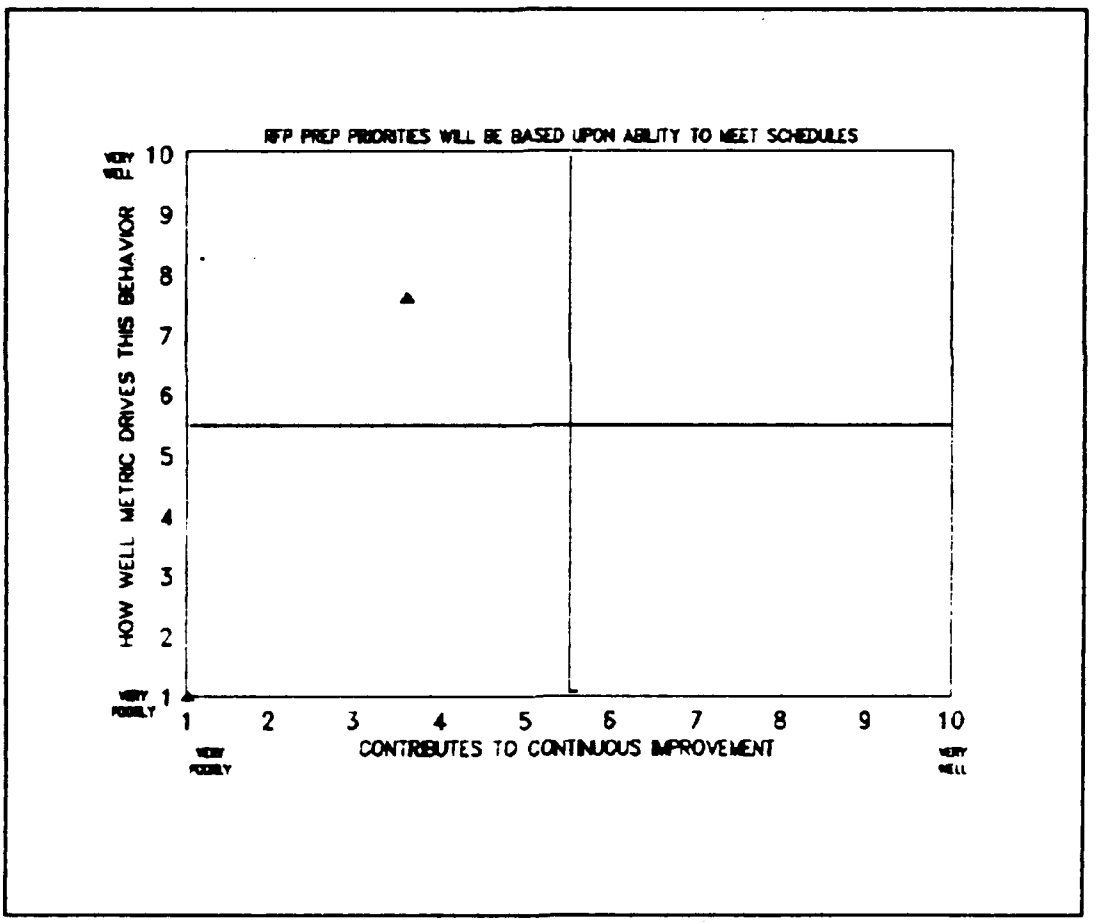


Figure A-67. Metric Five, Behavior Twelve

Behavior 13. Reduce the number of offices required to coordinate on package

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	2	-	-	-	-	-	3	-	-	5.60
How well does behavior contribute to CI?	1	1	-	2	-	-	-	1	-	-	3.80

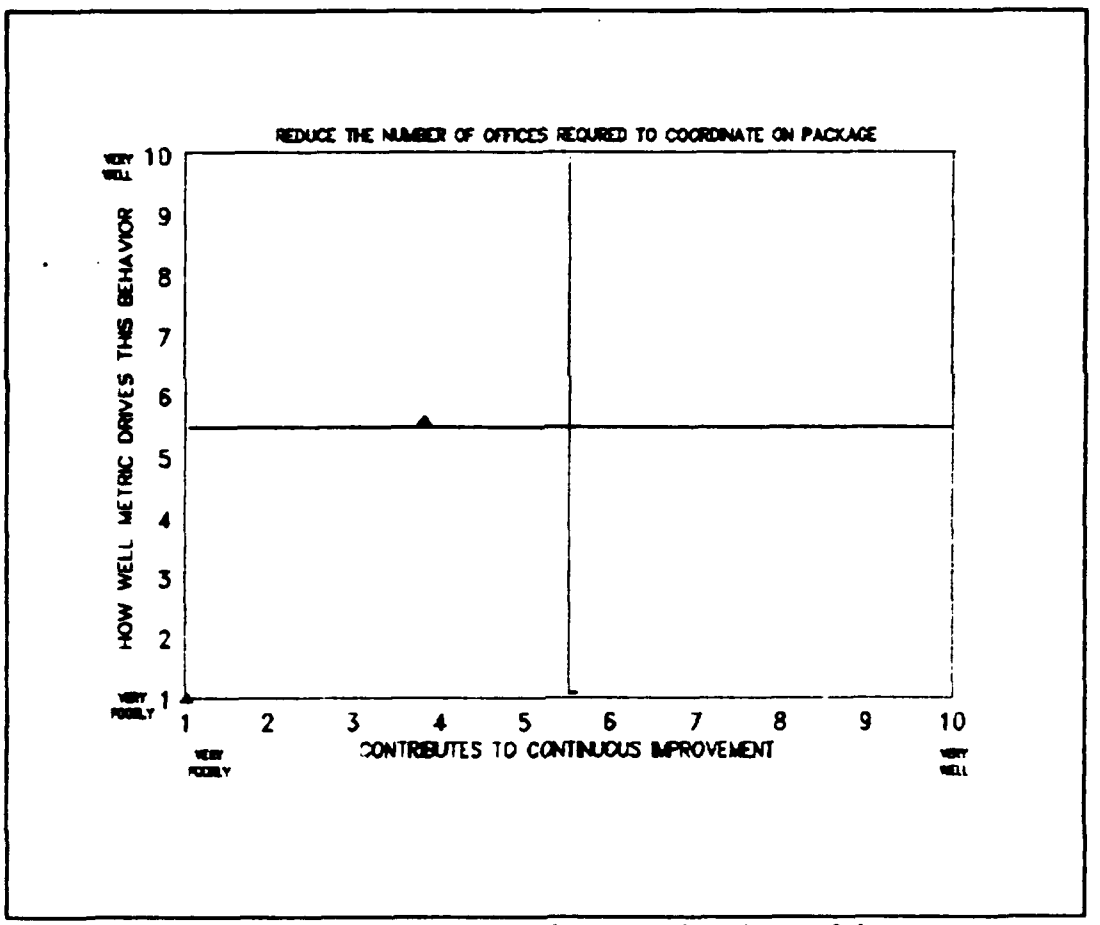


Figure A-68. Metric Five, Behavior Thirteen

Behavior 14. Have team comprised of functionals write package vs one functional organization

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	1	1	2	-	-	6.80
How well does behavior contribute to CI?	-	-	-	-	1	-	1	1	1	1	7.80

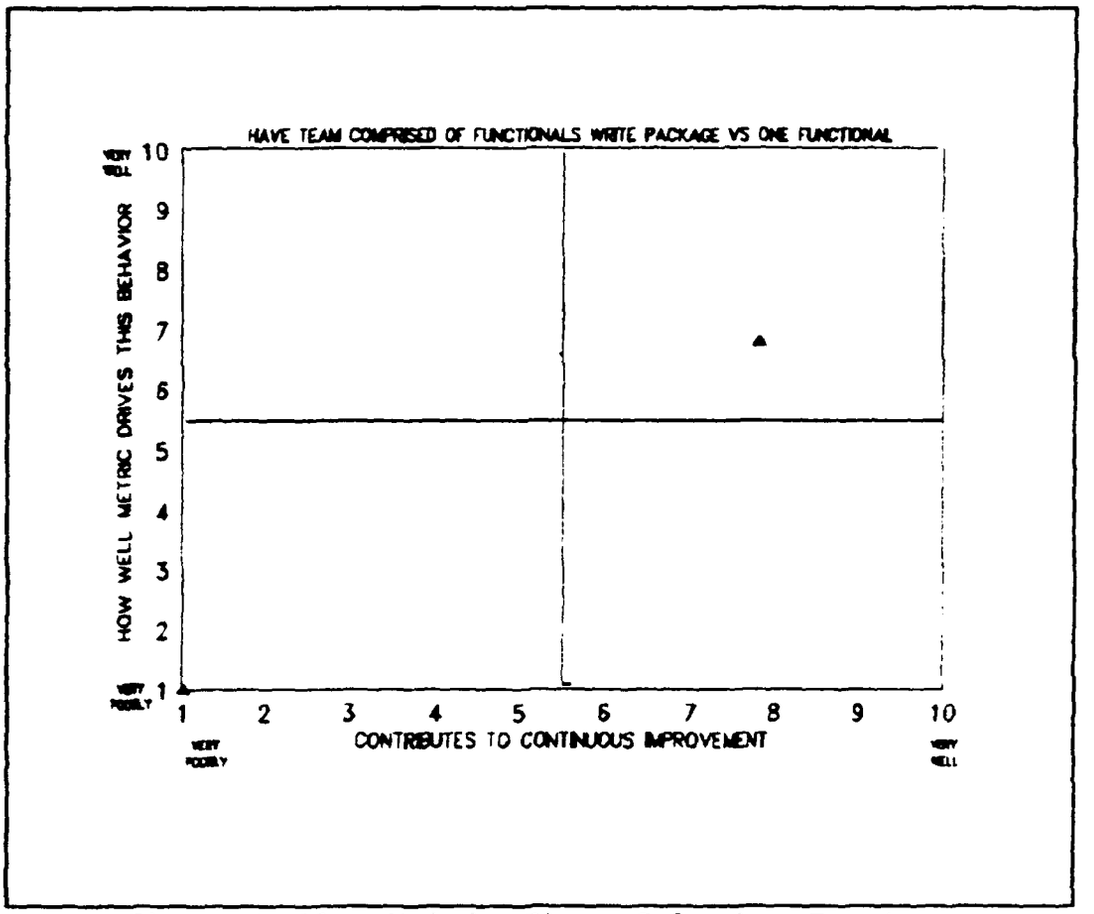


Figure A-69. Metric Five, Behavior Fourteen

Behavior 15. SPO will try to accelerate all RFP preps as much as possible

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	2	1	1	-	-	1	6.60
How well does behavior contribute to CI?	-	-	1	-	-	1	-	2	-	1	7.00

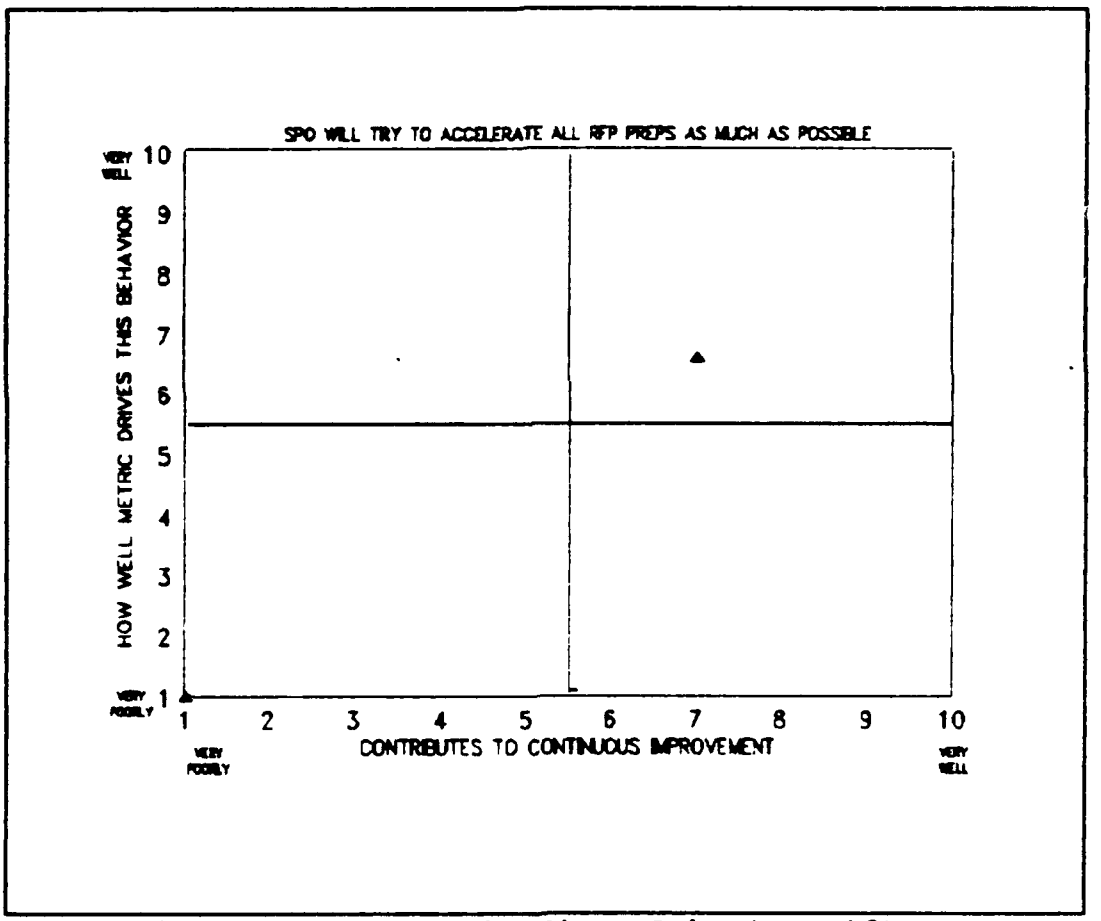


Figure A-70. Metric Five, Behavior Fifteen

Behavior 16. SPO will not wait on clear requirements definition from the user in order to meet schedule

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	1	1	1	1	-	-	5.80
How well does behavior contribute to CI?	1	1	1	1	1	-	-	-	-	-	3.00

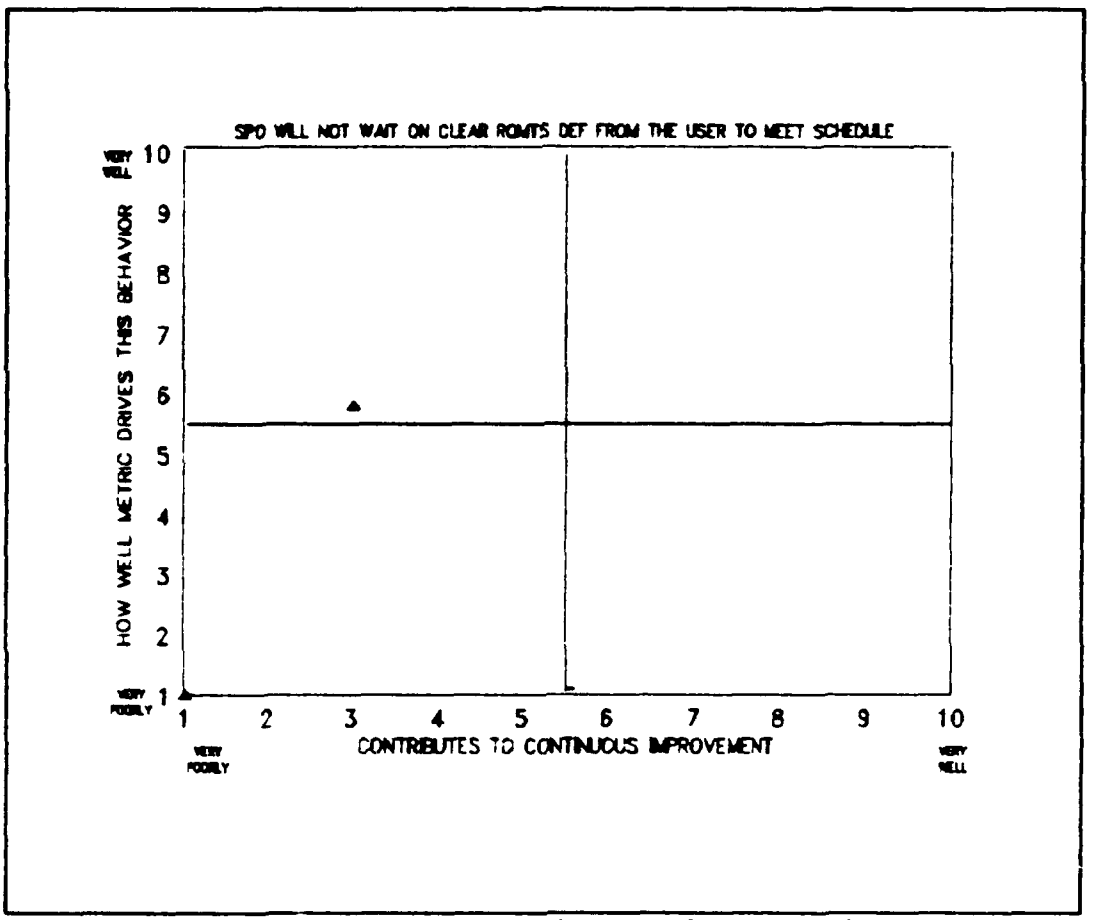


Figure A-71. Metric Five, Behavior Sixteen

Behavior 17. Schedule CCBs on as-needed date versus routine basis

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	2	1	1	1	-	7.20
How well does behavior contribute to CI?	-	-	-	-	2	1	1	-	1	-	6.40

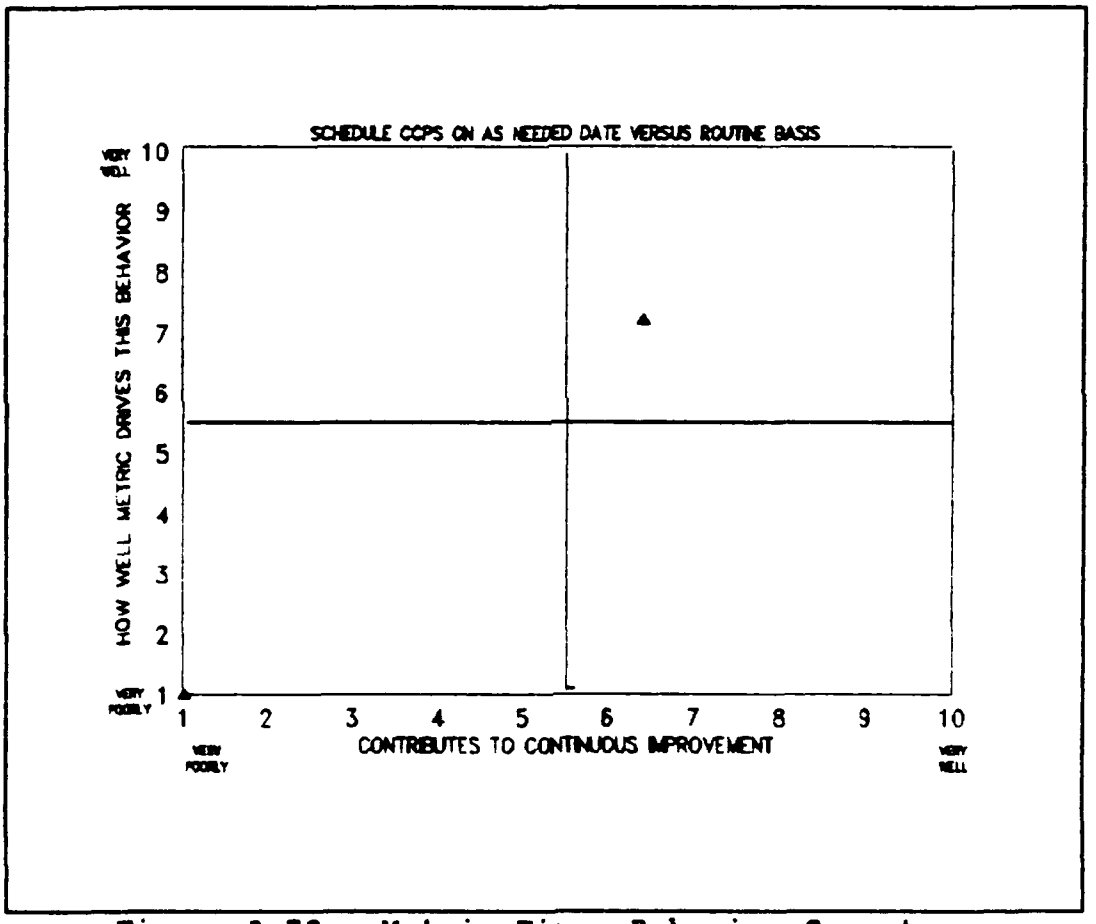


Figure A-72. Metric Five, Behavior Seventeen

Behavior 18. SPO will try to get informal jump on RFP prep before official approval

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	-	-	1	2	-	1	7.40
How well does behavior contribute to CI?	-	-	-	-	-	1	1	1	1	1	8.00

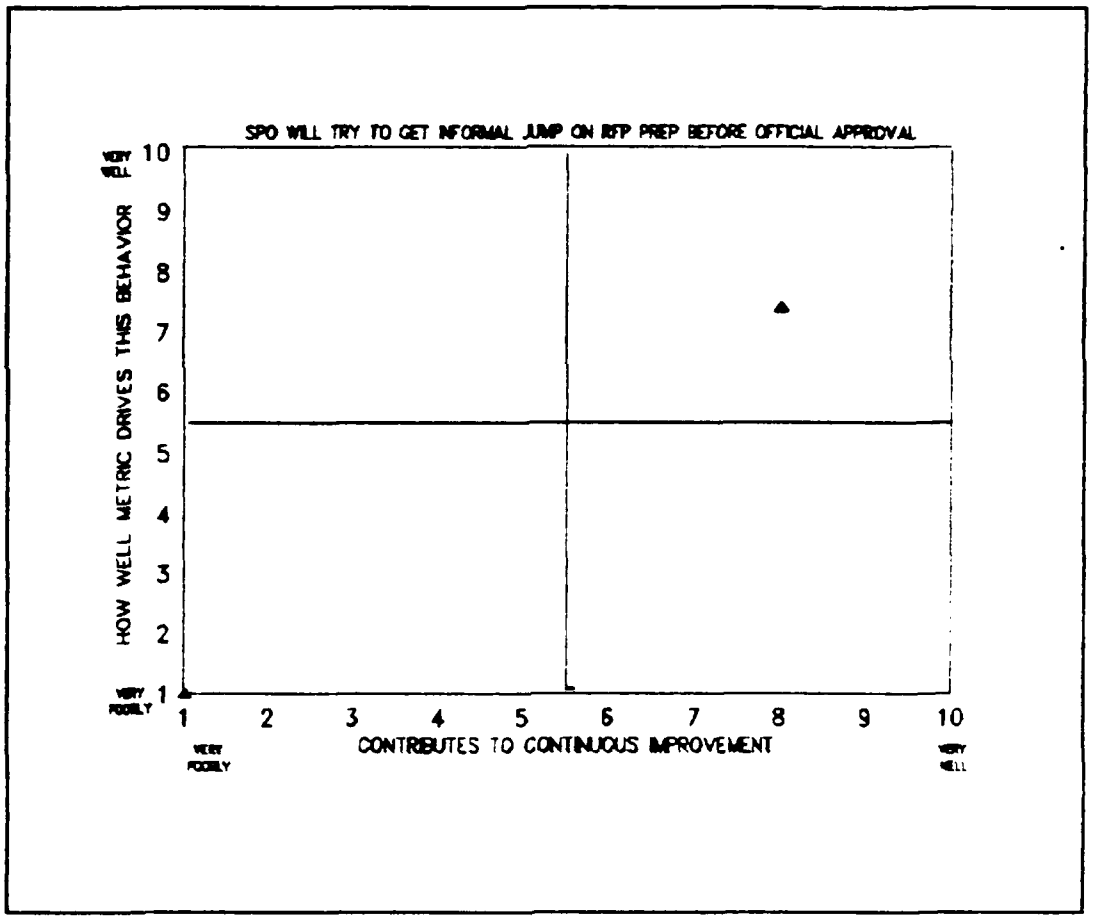


Figure A-73. Metric Five, Behavior Eighteen

Behavior 19. "Shotgun" proposals to each office vs routine coordination in-turn

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	3	1	-	1	7.80
How well does behavior contribute to CI?	-	-	-	-	2	-	2	-	-	1	6.80

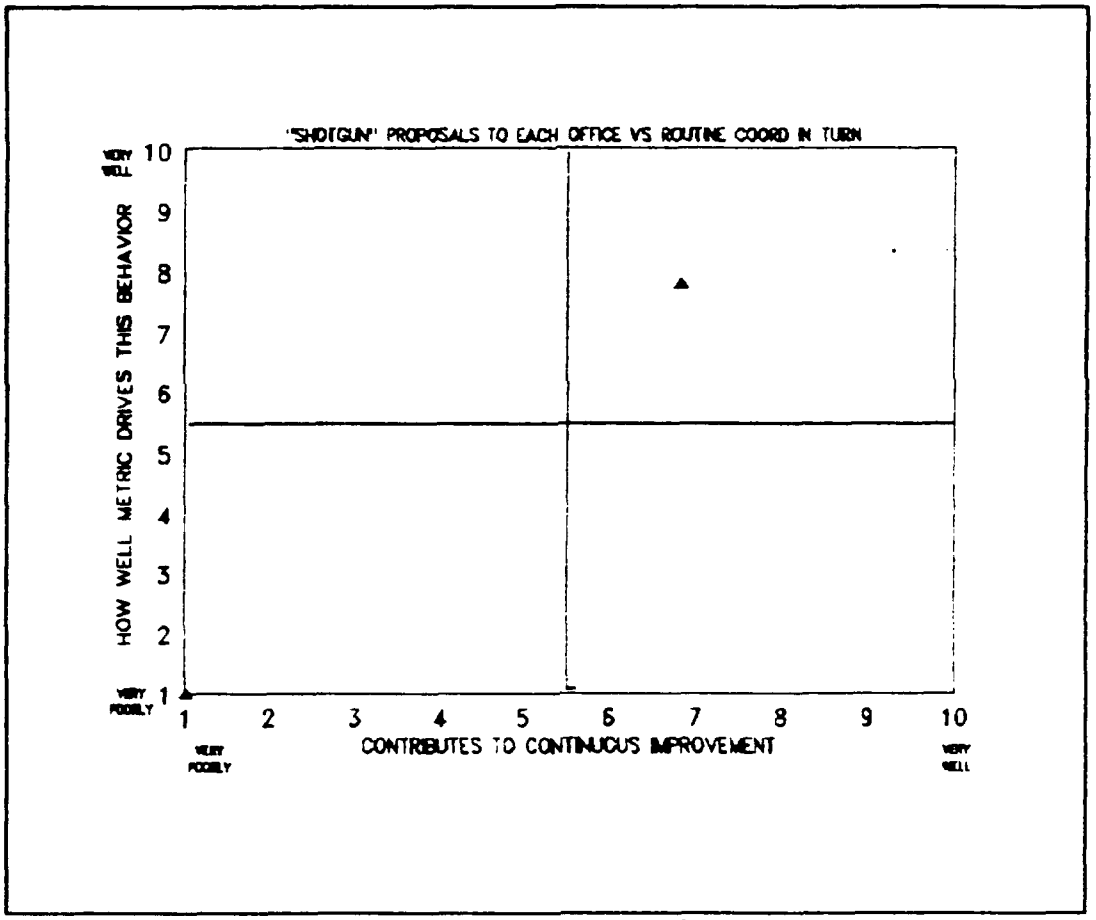


Figure A-74. Metric Five, Behavior Nineteen

Behavior 20. SPO will do a large number of small ECPs

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	2	-	1	1	-	-	5.80
How well does behavior contribute to CI?	-	1	1	1	1	1	-	-	-	-	4.00

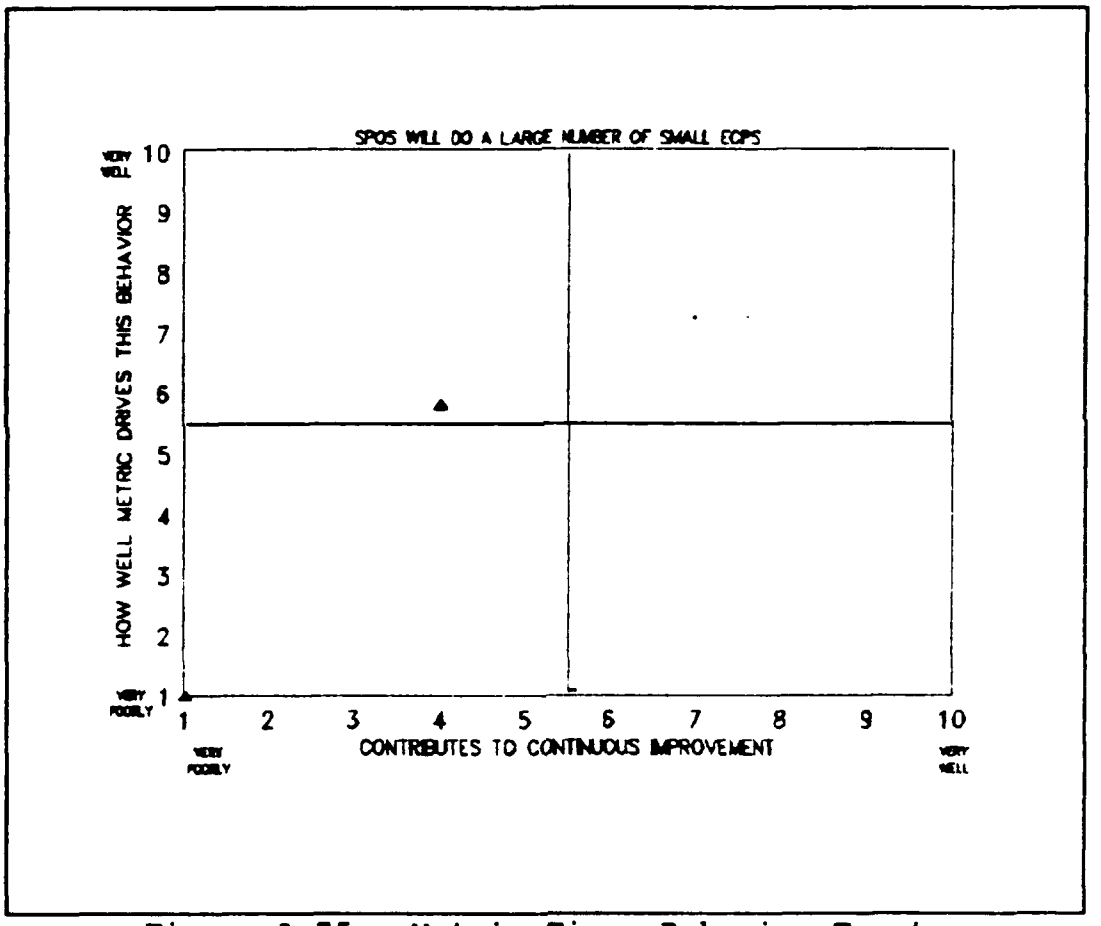


Figure A-75. Metric Five, Behavior Twenty

Behavior 21. The SPO will be less likely to take on multiple RFPs simultaneously

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	2	-	-	2	-	1	-	6.20
How well does behavior contribute to CI?	-	-	-	2	1	2	-	-	-	-	5.00

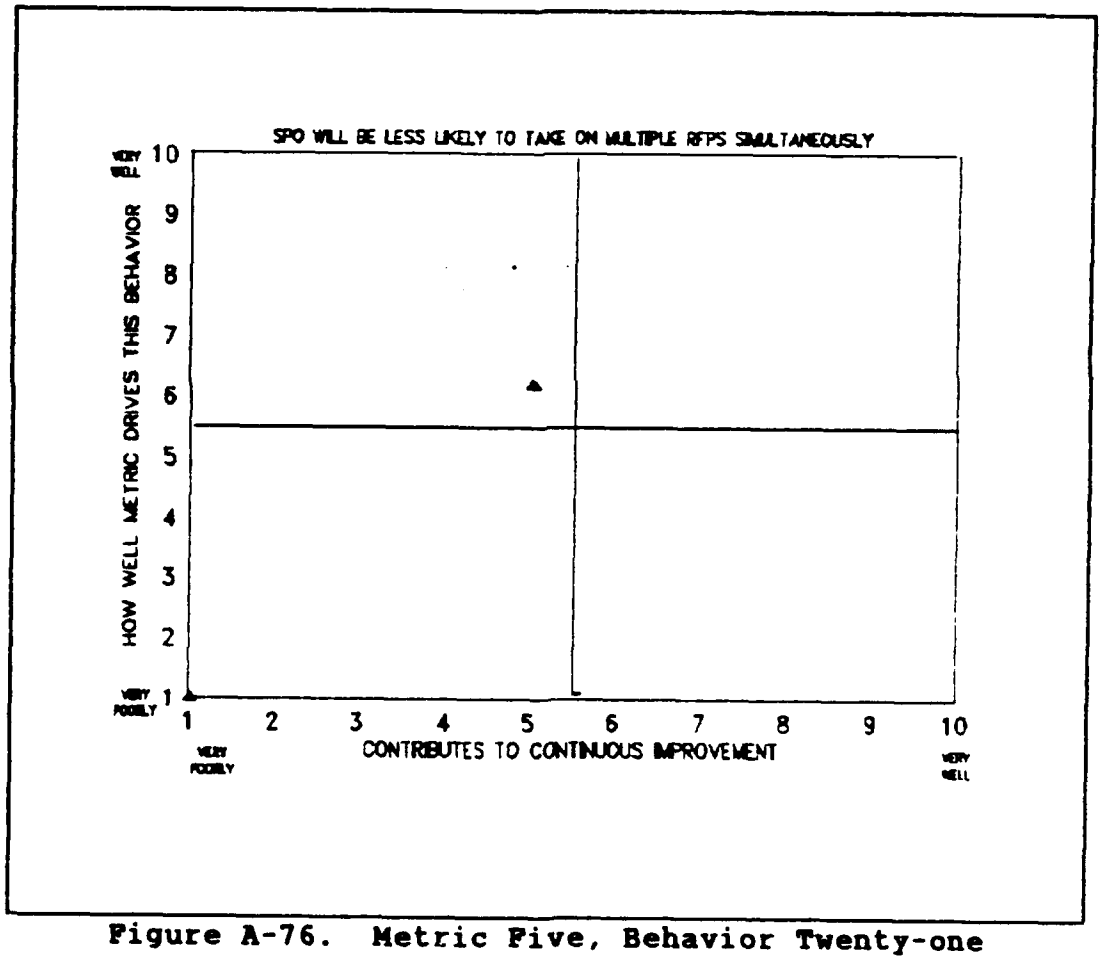


Figure A-76. Metric Five, Behavior Twenty-one

Metric 6 Behavior 1. SPO EDMO will be more concerned with data submittal dates than with data content

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	2	-	-	-	-	-	-	2	-	1	5.60
How well does behavior contribute to CI?	-	3	-	1	-	-	-	-	-	1	4.00

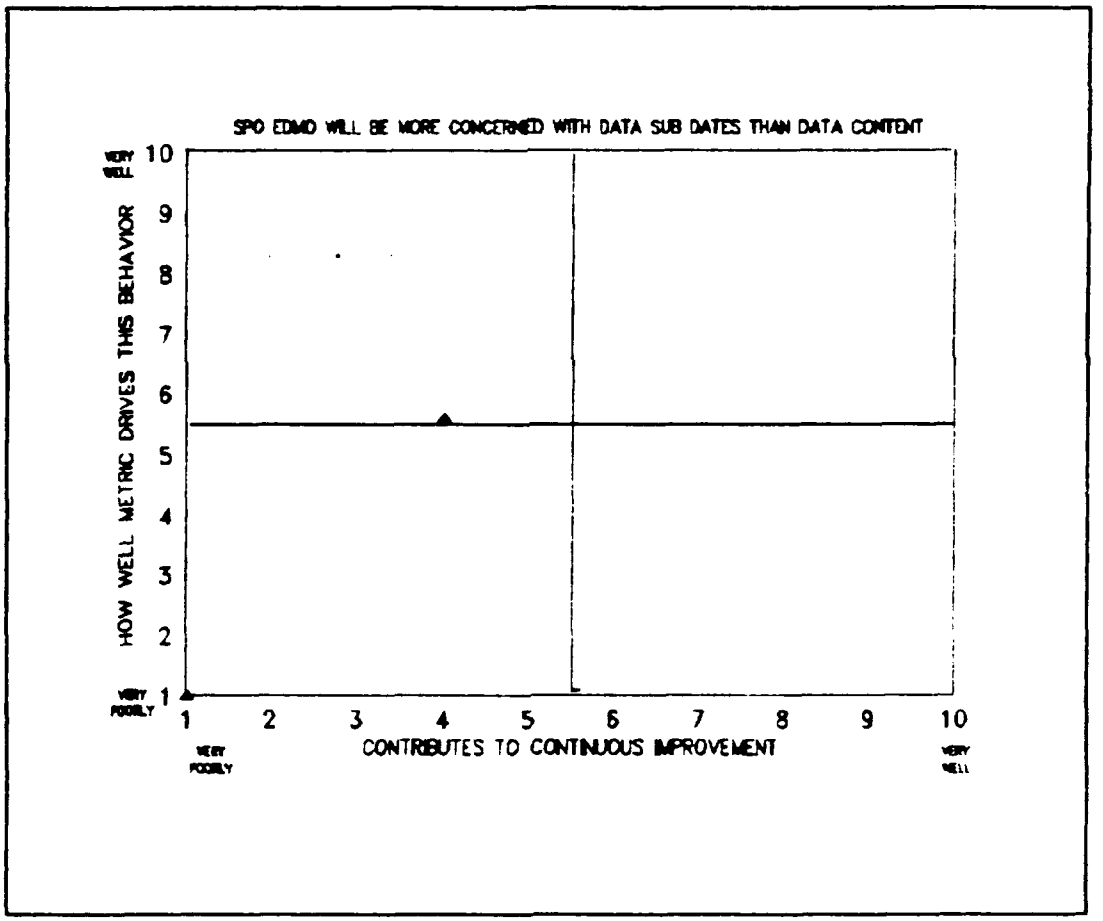


Figure A-77. Metric Six, Behavior One

Behavior 2. SPO will focus on the contractor versus SPO review/approval

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	1	-	-	-	-	1	-	3	-	-	6.20
How well does behavior contribute to CI?	1	1	-	1	-	-	-	2	-	-	4.60

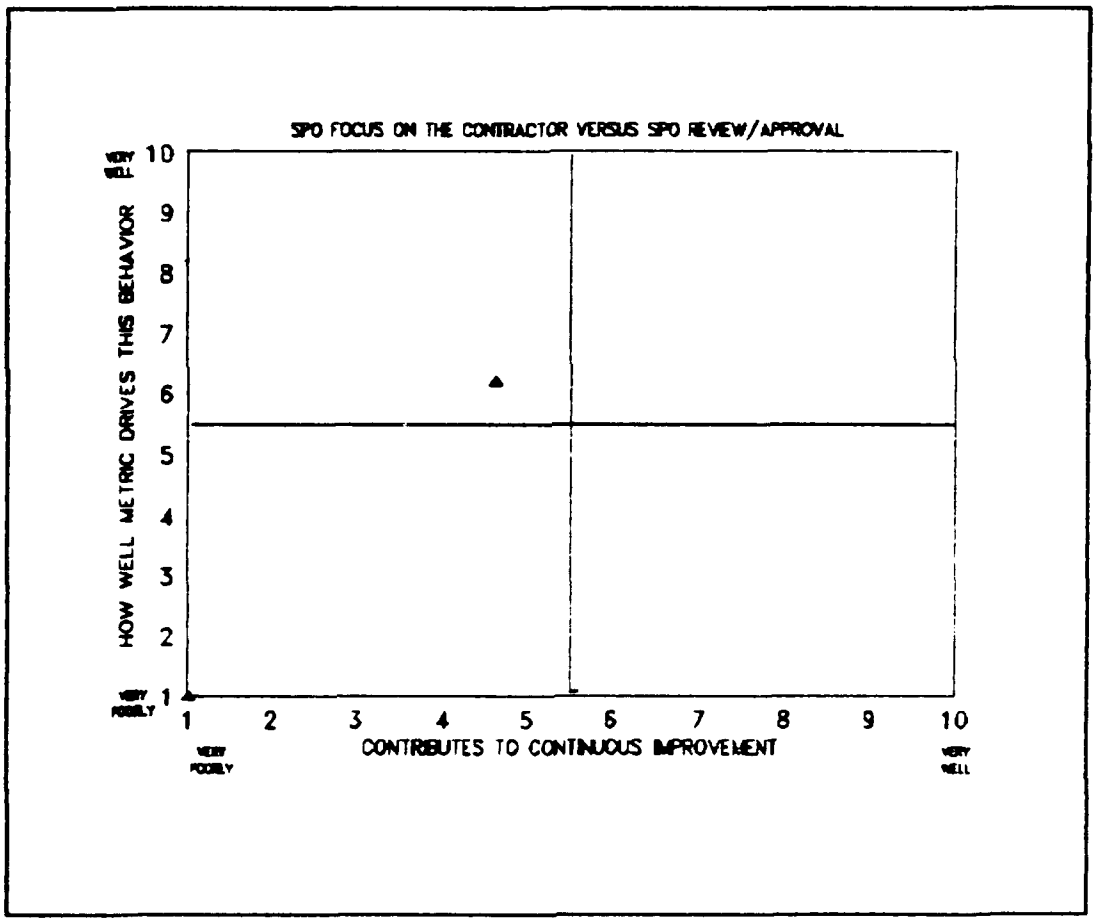


Figure A-78. Metric Six, Behavior Two

Behavior 3. Contractor will be more concerned with data submittal dates than with data content

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	-	3	-	2	8.80
How well does behavior contribute to CI?	2	2	-	1	-	-	-	-	-	-	2.00

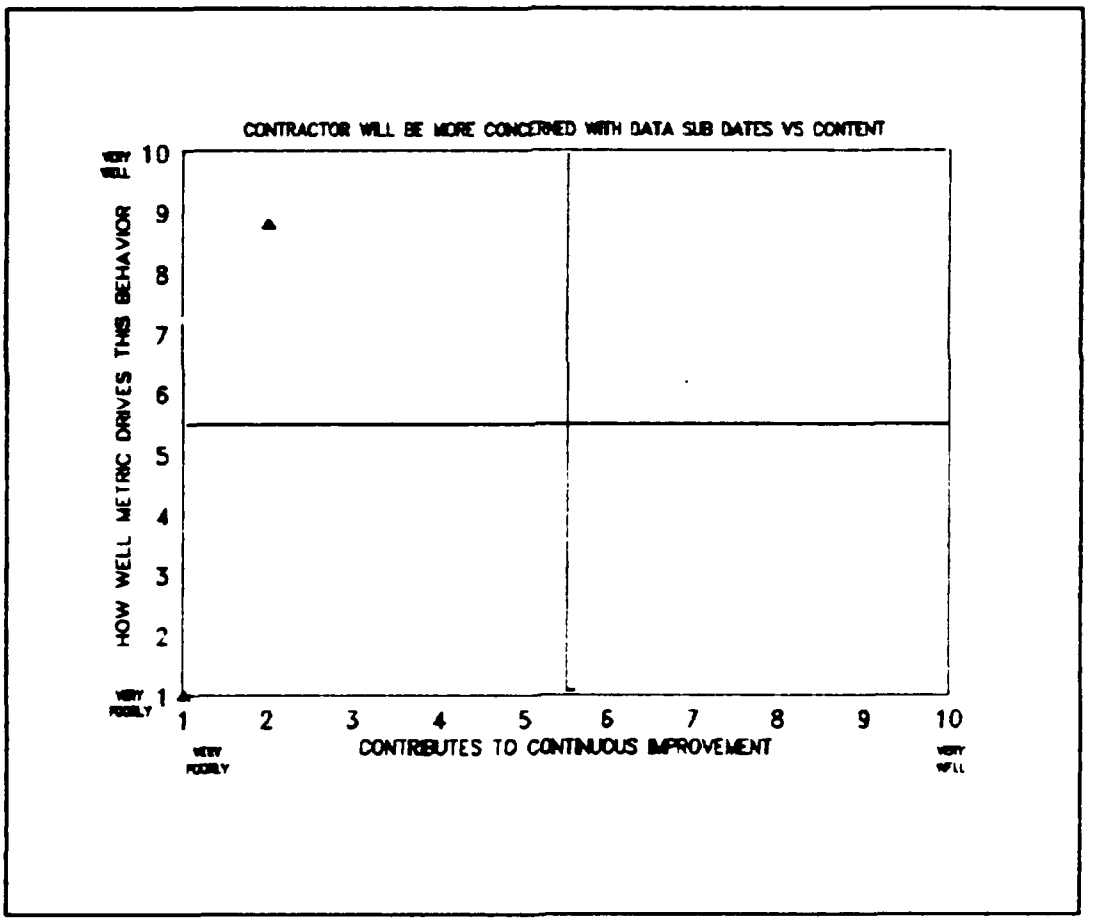


Figure A-79. Metric Six, Behavior Three

Behavior 4. Weekly status reports generated and tracked by SPO data manager

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	1	-	-	-	1	-	1	1	1	7.00
How well does behavior contribute to CI?	-	-	2	-	-	1	-	2	-	-	5.60

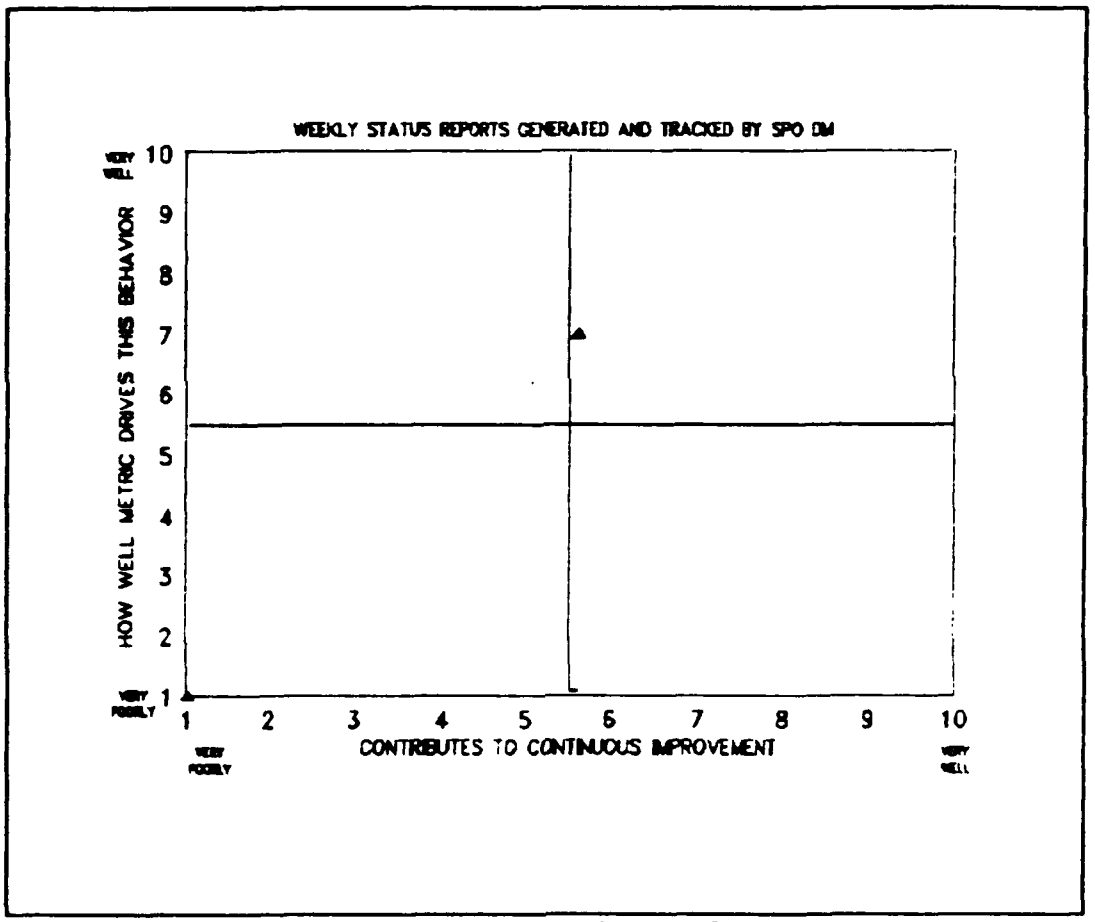


Figure A-80. Metric Six, Behavior Four

Behavior 5. Contractor will request extensions of due dates to later in program

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	2	-	1	1	1	-	6.80
How well does behavior contribute to CI?	1	-	-	2	2	-	-	-	-	-	3.80

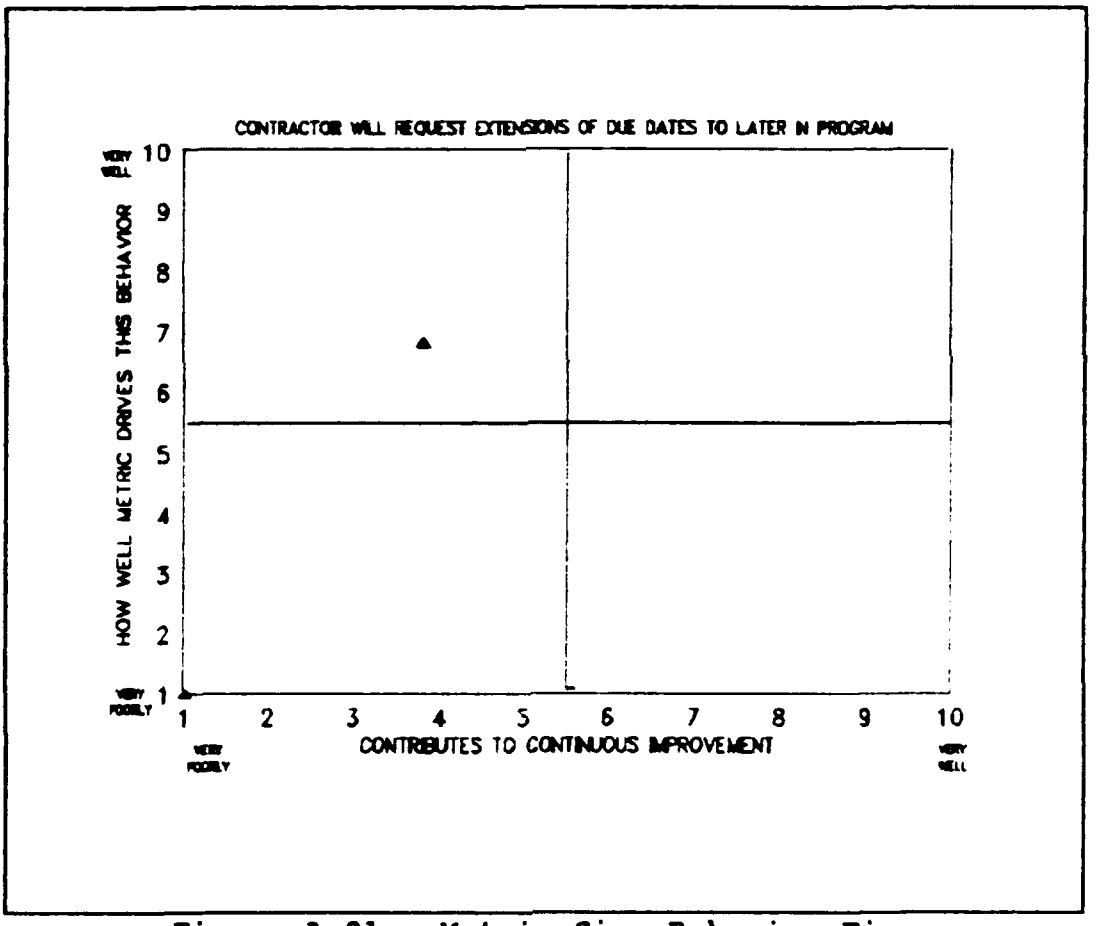


Figure A-81. Metric Six, Behavior Five

Behavior 6. SPO personnel will not review data; they'll just track its submittal

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	3	-	-	1	-	-	1	-	-	-	2.80
How well does behavior contribute to CI?	2	2	-	-	1	-	-	-	-	-	2.20

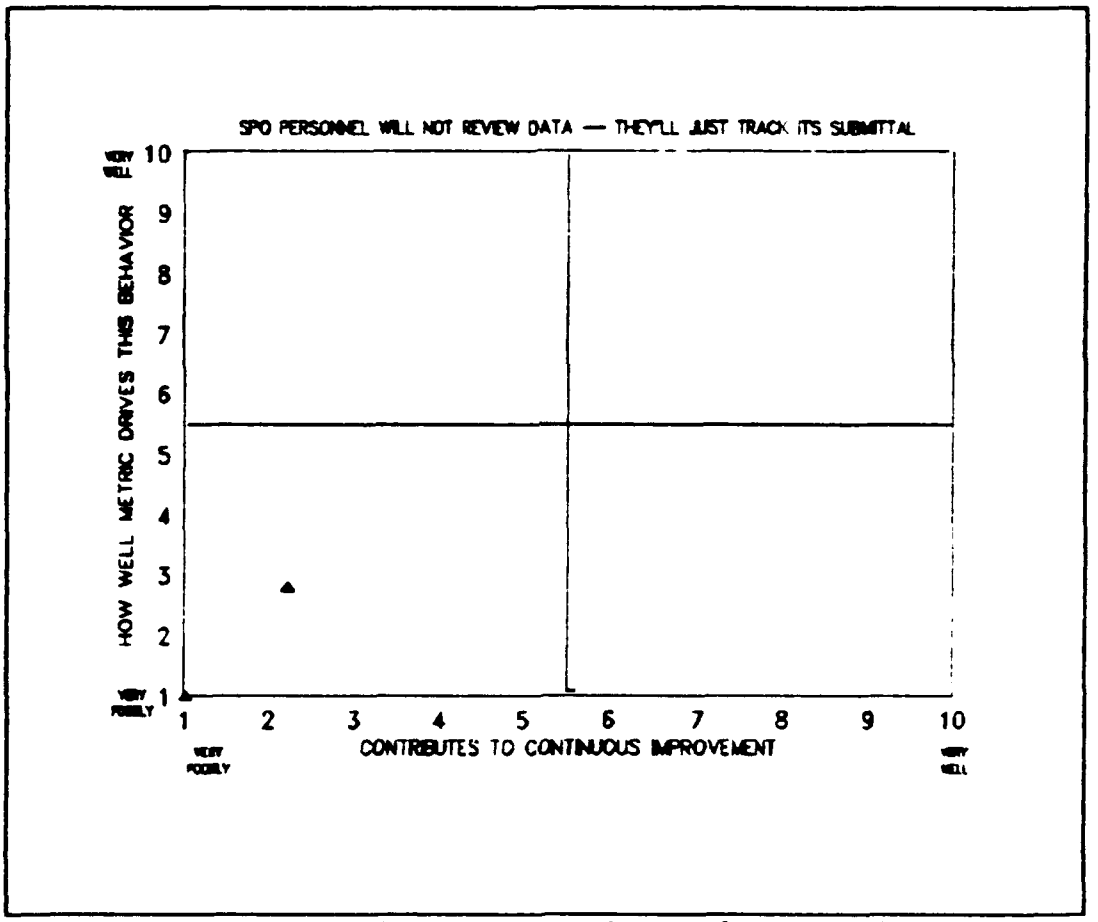


Figure A-82. Metric Six, Behavior Six

Behavior 7. Contractor will request SPO to reduce data requirements

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	-	-	1	2	1	-	7.20
How well does behavior contribute to CI?	1	-	-	1	2	-	-	1	-	-	4.60

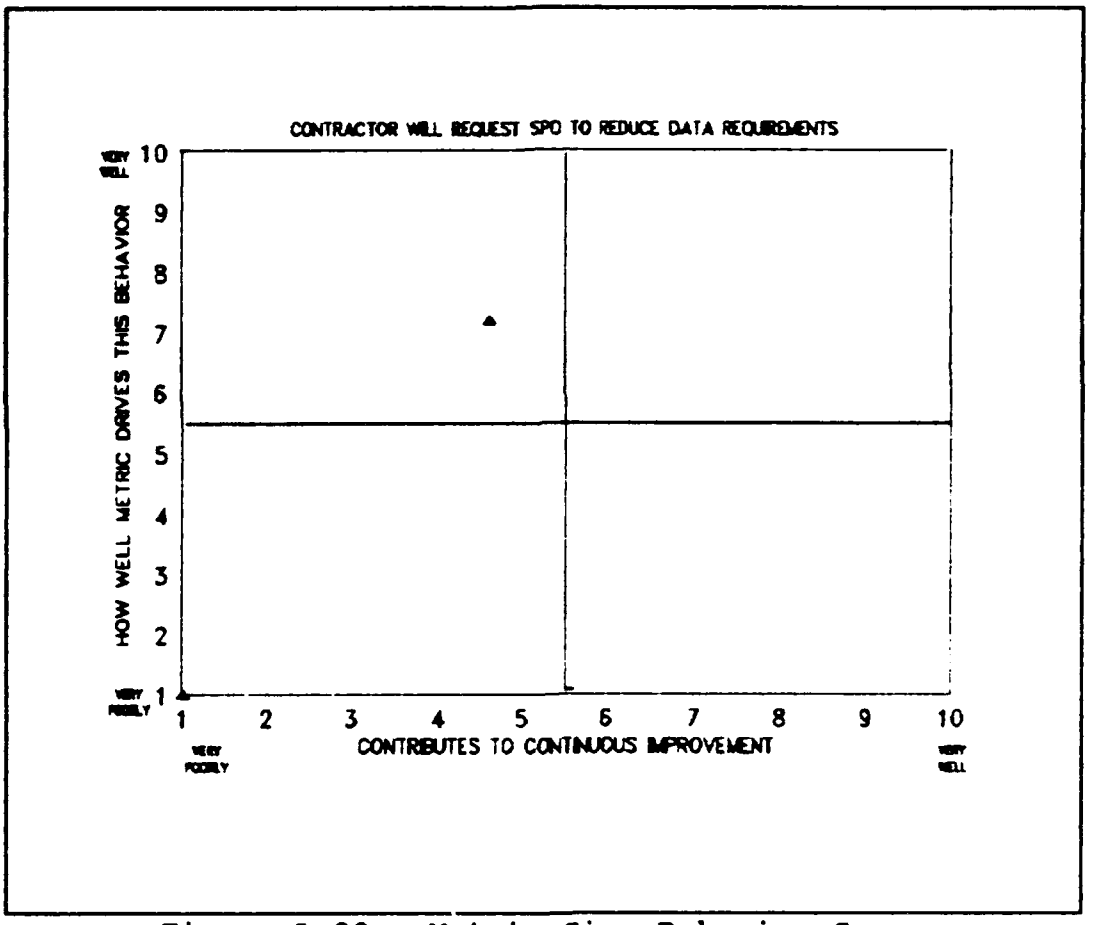


Figure A-83. Metric Six, Behavior Seven

Behavior 8. SPO will conduct periodic review of essential data requirements for additions/deletions

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	2	-	-	2	-	1	-	-	5.20
How well does behavior contribute to CI?	-	-	-	-	1	-	-	3	1	-	7.60

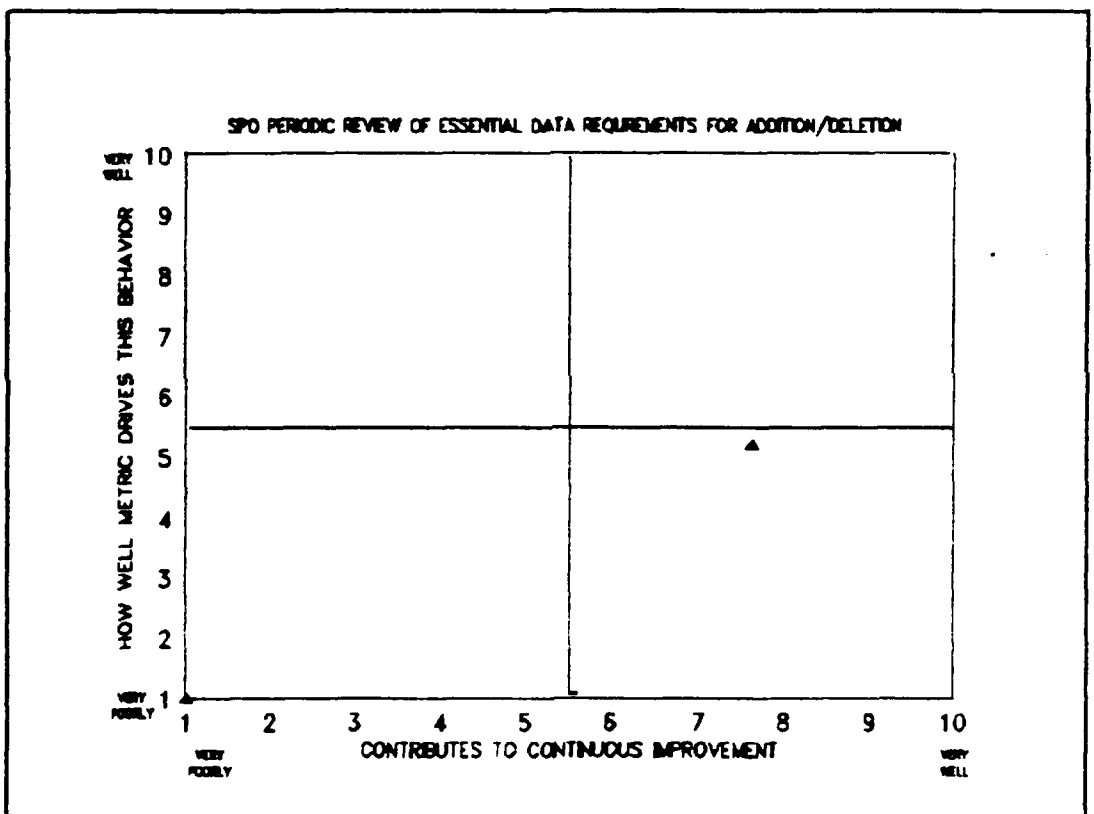


Figure A-84. Metric Six, Behavior Eight

Behavior 9. SPO does in-depth tracking of the contractor data management system

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	1	-	-	-	1	2	-	1	-	6.20
How well does behavior contribute to CI?	-	-	2	-	-	-	-	1	1	1	6.60

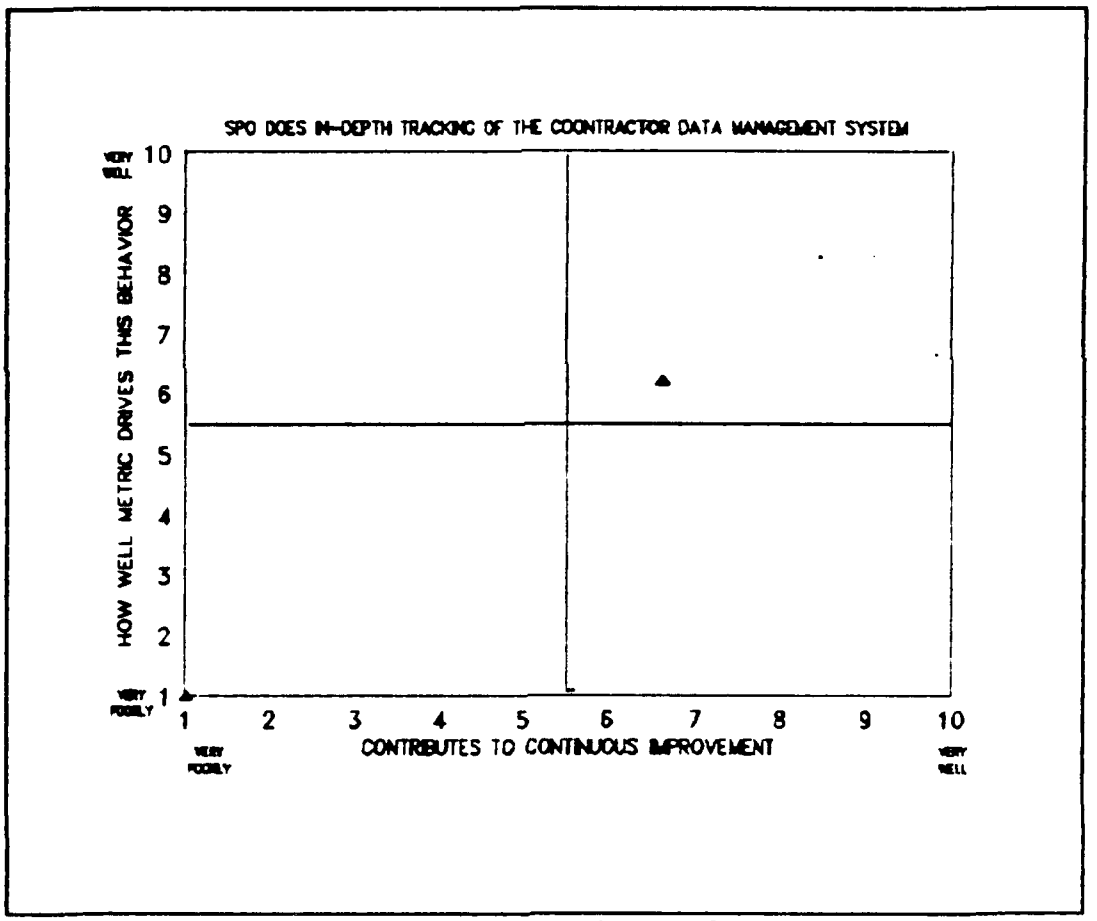


Figure A-85. Metric Six, Behavior Nine

Behavior 10. Contractor will shift manpower to work on late data submittals

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	1	3	-	1	8.20
How well does behavior contribute to CI?	-	-	-	1	2	-	-	2	-	-	6.00

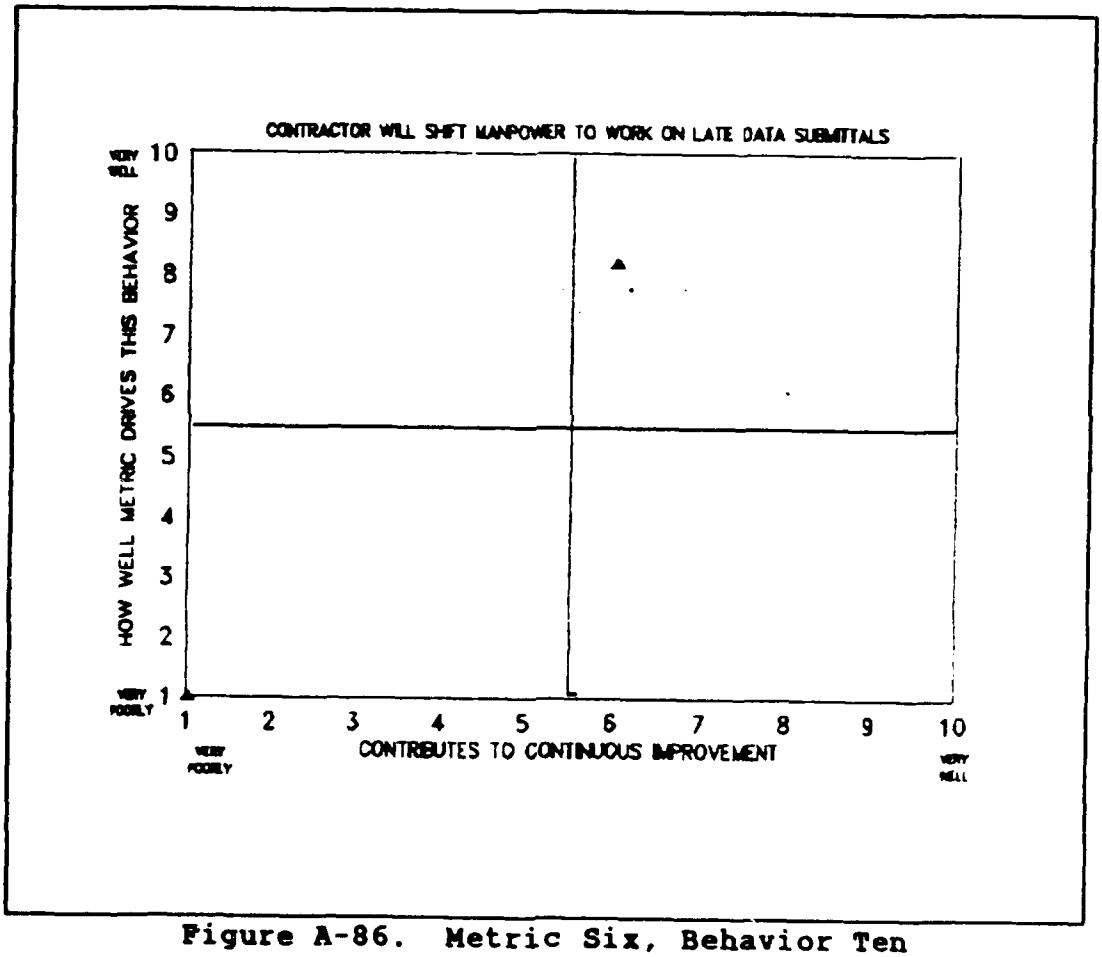


Figure A-86. Metric Six, Behavior Ten

Behavior 11. Contractor will find excuses for late data deliveries

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	2	1	1	1	8.20
How well does behavior contribute to CI?	1	3	1	-	-	-	-	-	-	-	2.00

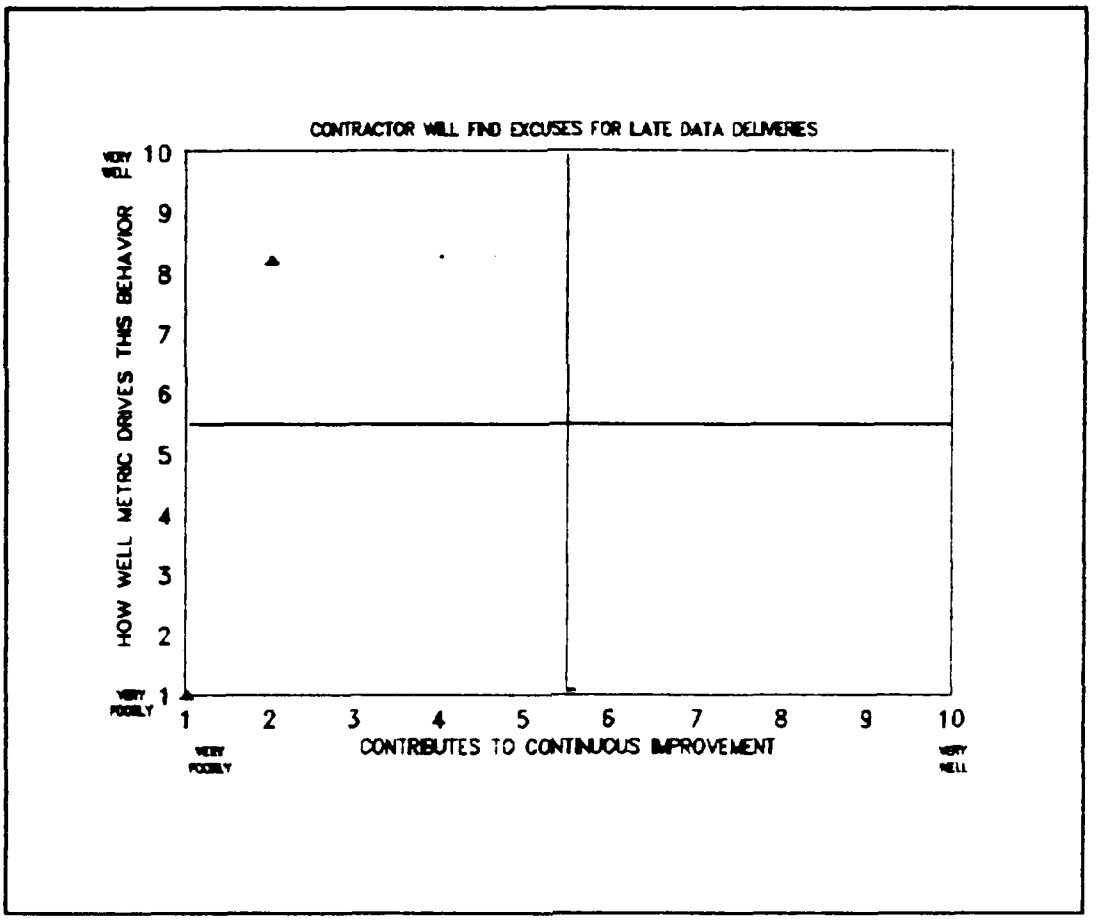


Figure A-87. Metric Six, Behavior Eleven

Behavior 12. SPO will seek to do up front work with contractor to improve submittals

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	2	-	-	-	1	1	1	-	6.00
How well does behavior contribute to CI?	-	-	-	-	-	-	-	2	2	1	8.80

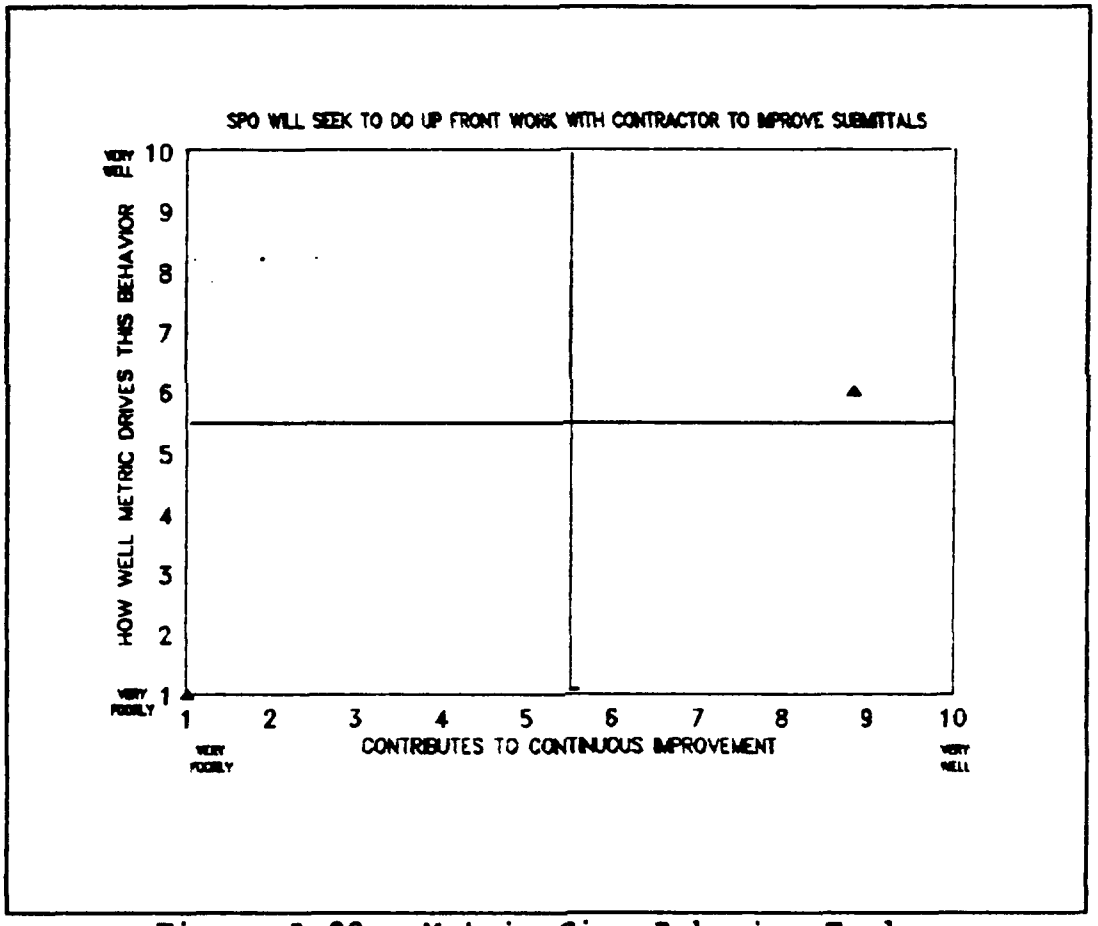


Figure A-88. Metric Six, Behavior Twelve

Behavior 13. The status of data submissions will be briefed at program reviews

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	2	1	2	-	8.00
How well does behavior contribute to CI?	1	-	-	3	-	-	-	1	-	-	4.20

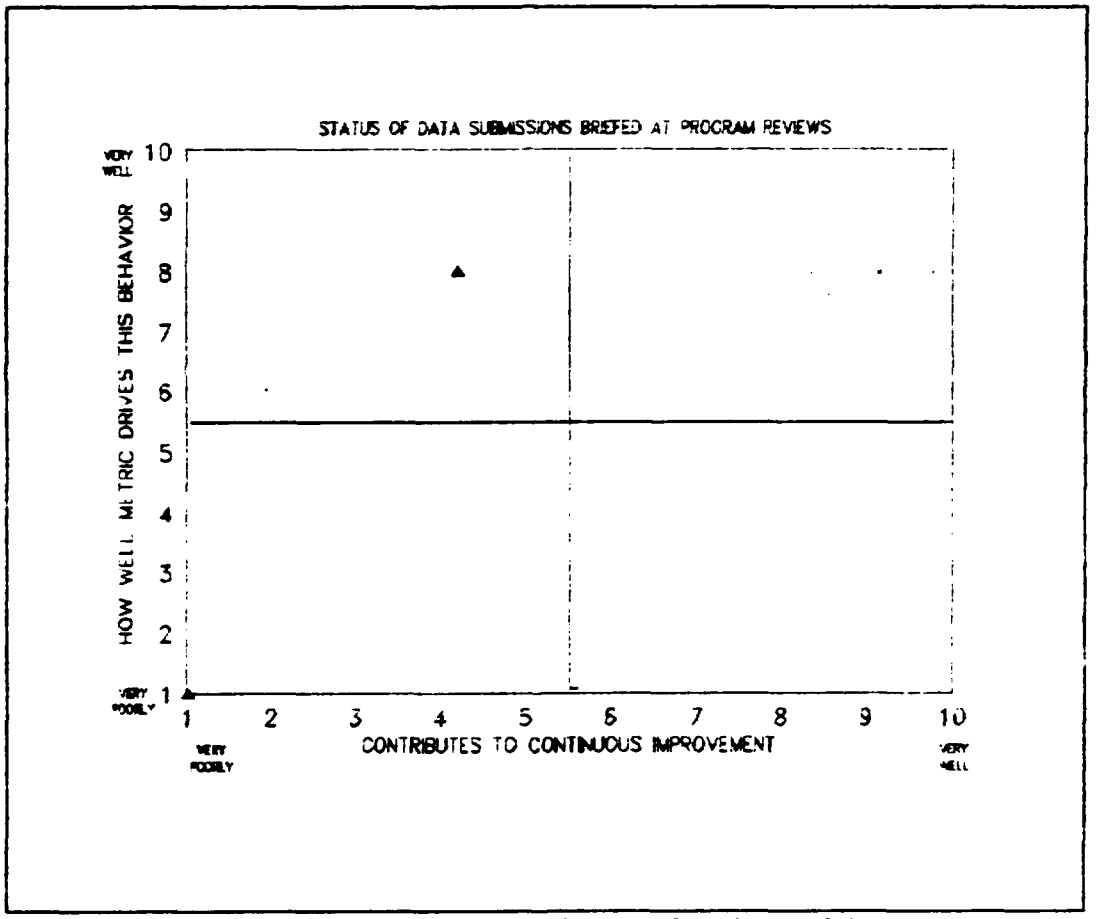


Figure A-89. Metric Six, Behavior Thirteen

Behavior 14. Contractor will develop a process to get quality data out in a timely manner

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	-	-	-	3	1	-	7.20
How well does behavior contribute to CI?	-	-	-	-	-	-	-	-	1	4	9.80

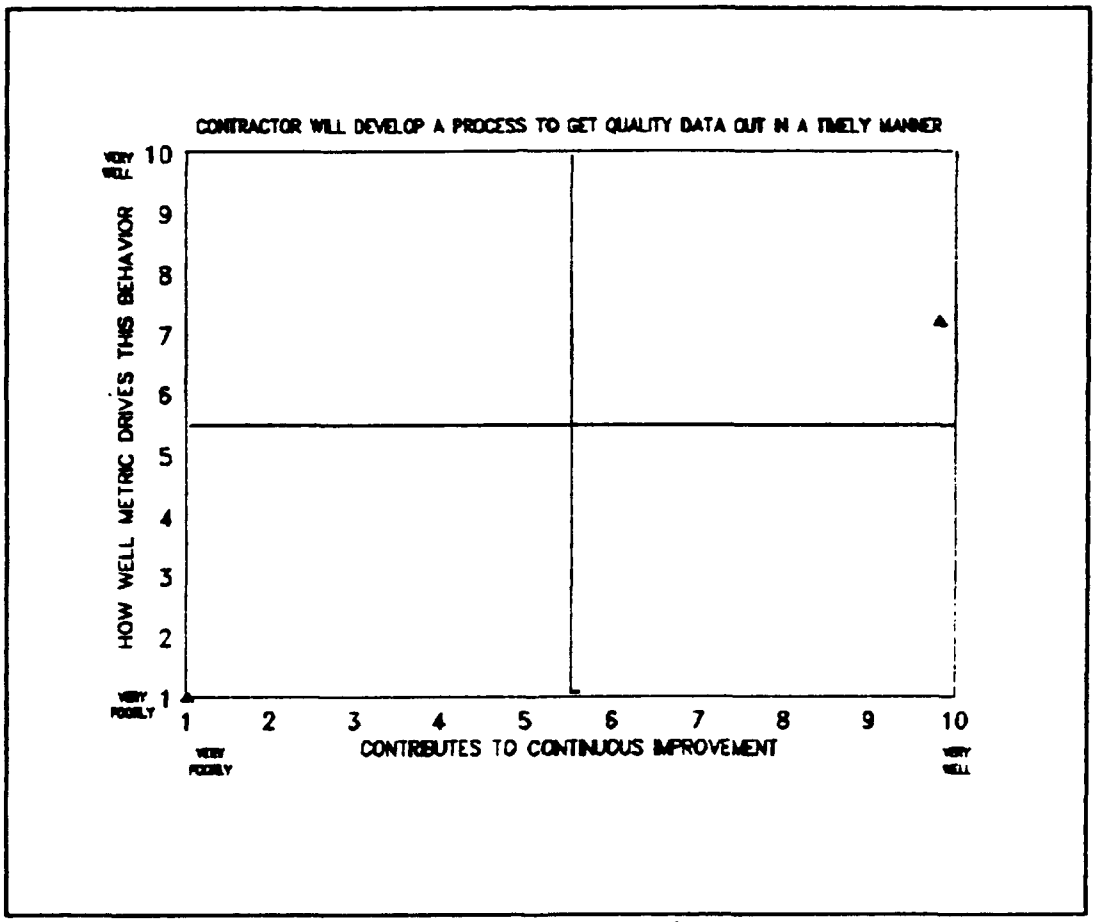


Figure A-90. Metric Six, Behavior Fourteen

Metric 7 Behavior 1. SPO increases review of GFE delivery status

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	-	1	2	-	2	8.60
How well does behavior contribute to CI?	-	-	-	1	1	-	1	-	-	2	7.20

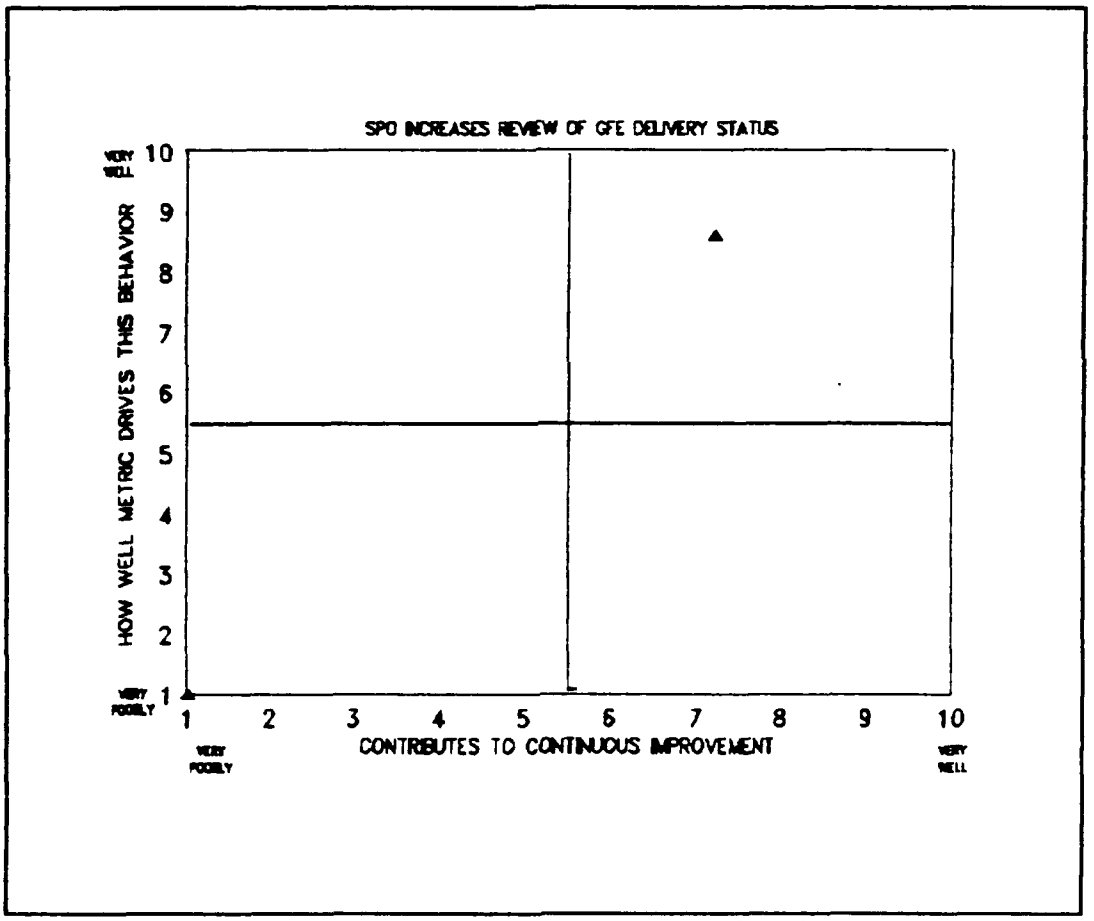


Figure A-91. Metric Seven, Behavior One

Behavior 2. SPO continually scrubs contractor stated requirements

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	-	3	1	-	7.60
How well does behavior contribute to CI?	-	-	-	-	-	-	1	1	1	2	8.80

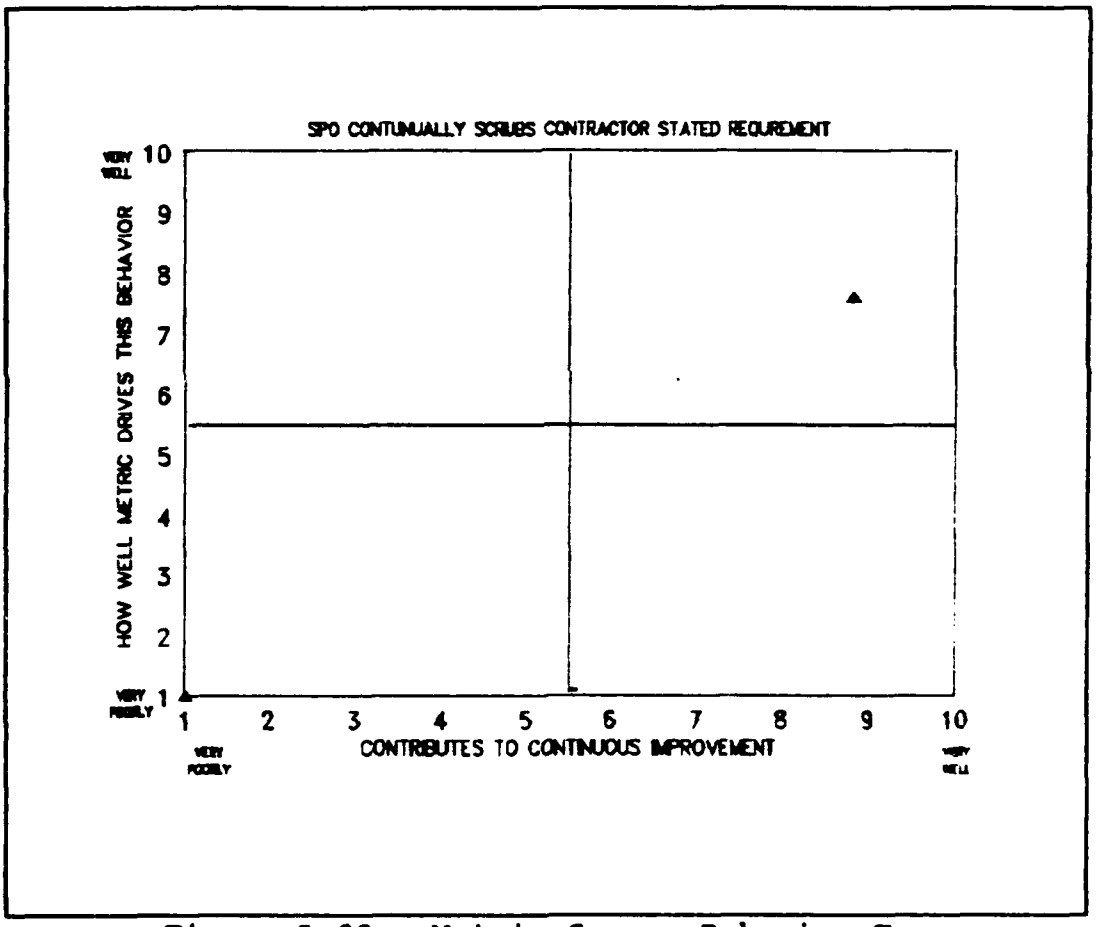


Figure A-92. Metric Seven, Behavior Two

Behavior 3. Government will work to match actual with required

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	1	1	1	1	8.00
How well does behavior contribute to CI?	-	-	-	-	-	-	2	-	1	2	8.60

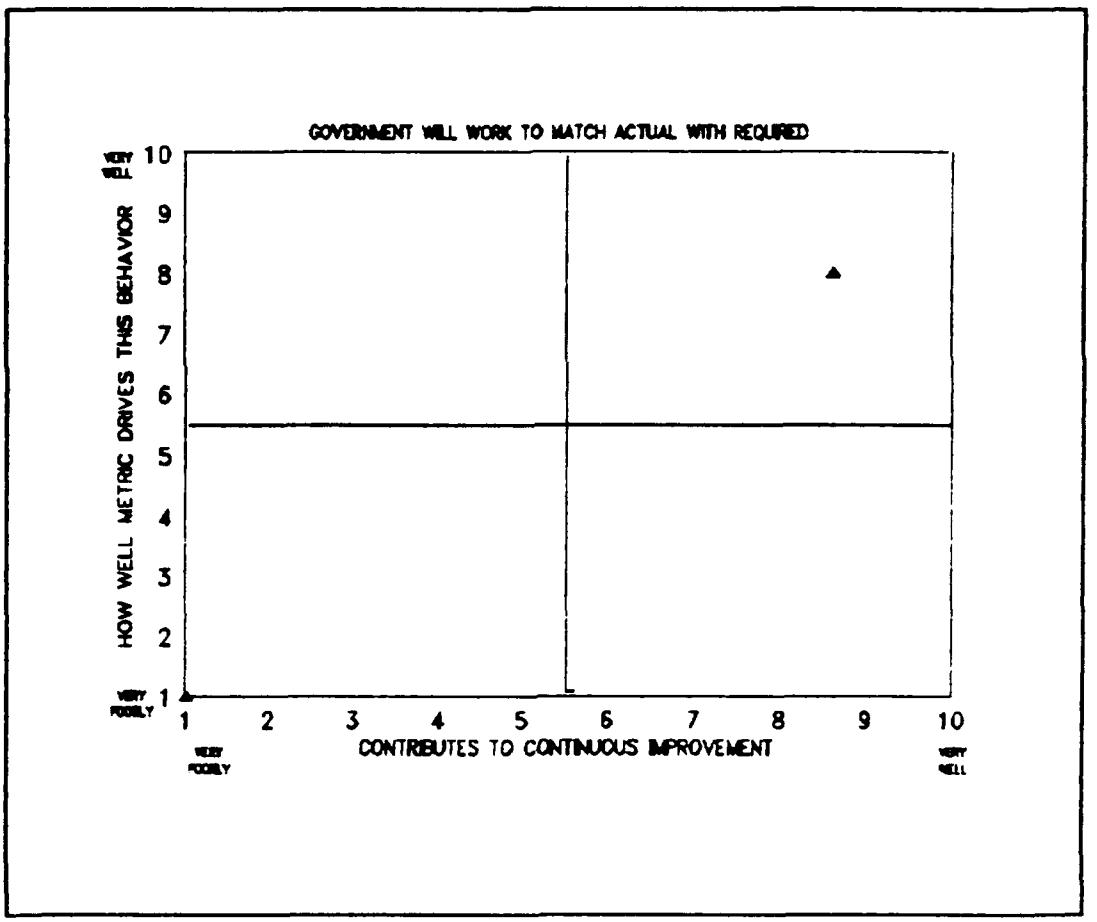


Figure A-93. Metric Seven, Behavior Three

Behavior 4. GFE deliveries will be prioritized by availability vice importance

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	-	1	1	2	-	-	6.60
How well does behavior contribute to CI?	-	-	3	2	-	-	-	-	-	-	3.40

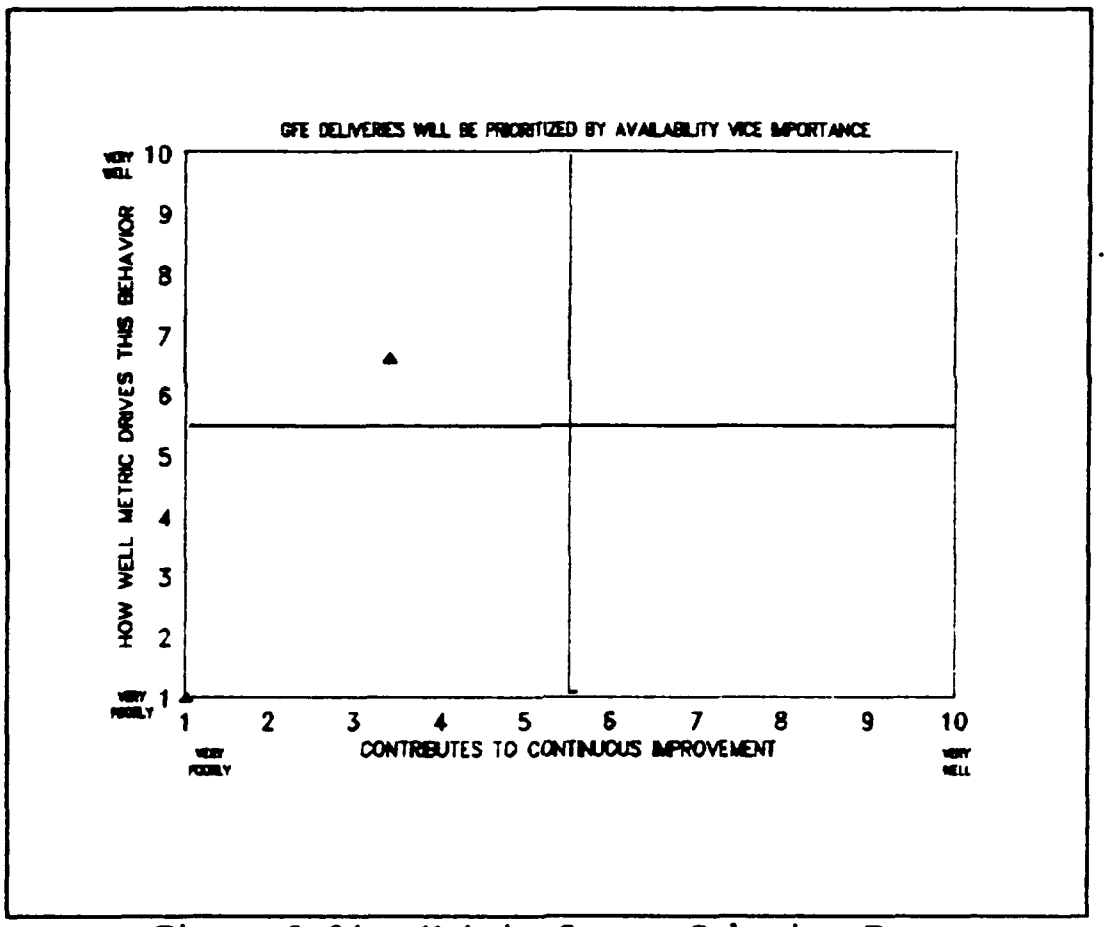


Figure A-94. Metric Seven, Behavior Four

Behavior 5. IM will pressure SPO to reduce GFE requirements

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	1	-	1	2	-	-	6.20
How well does behavior contribute to CI?	1	-	1	-	-	2	-	1	-	-	4.80

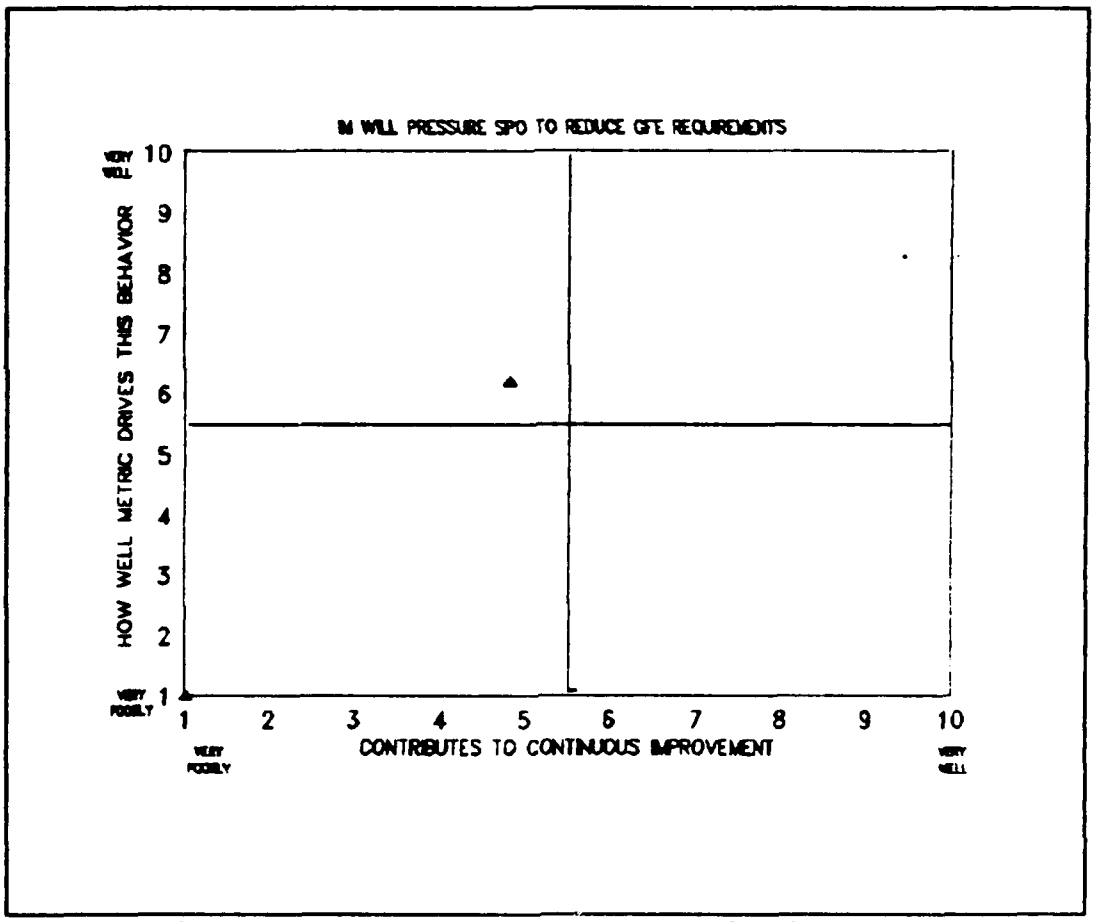


Figure A-95. Metric Seven, Behavior Five

Behavior 6. IM will attempt to increase GFE stocks

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	2	2	-	-	7.00
How well does behavior contribute to CI?	-	-	1	1	-	-	-	2	-	1	6.60

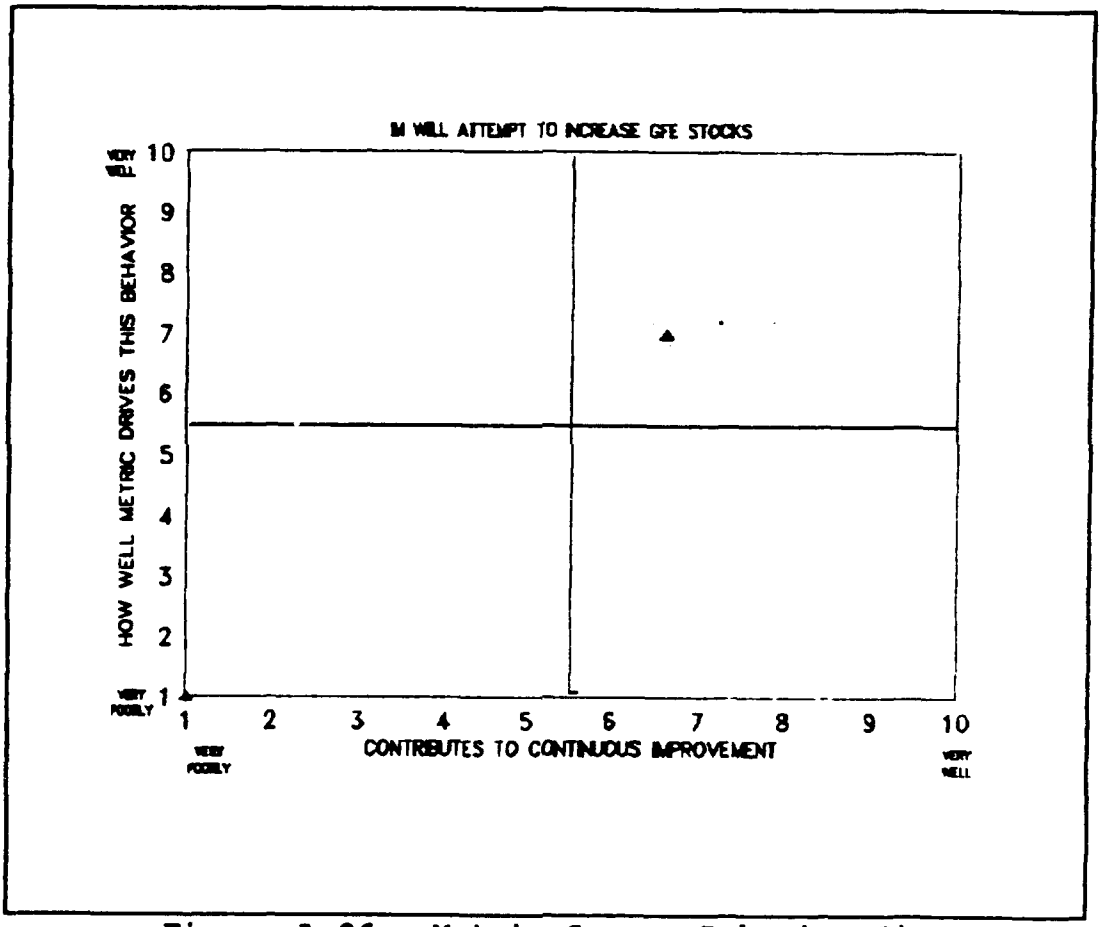


Figure A-96. Metric Seven, Behavior Six

Behavior 7. SPO acceptance of poor quality GFE in order to avoid delivery delays

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	3	1	-	-	7.00
How well does behavior contribute to CI?	2	3	-	-	-	-	-	-	-	-	1.60

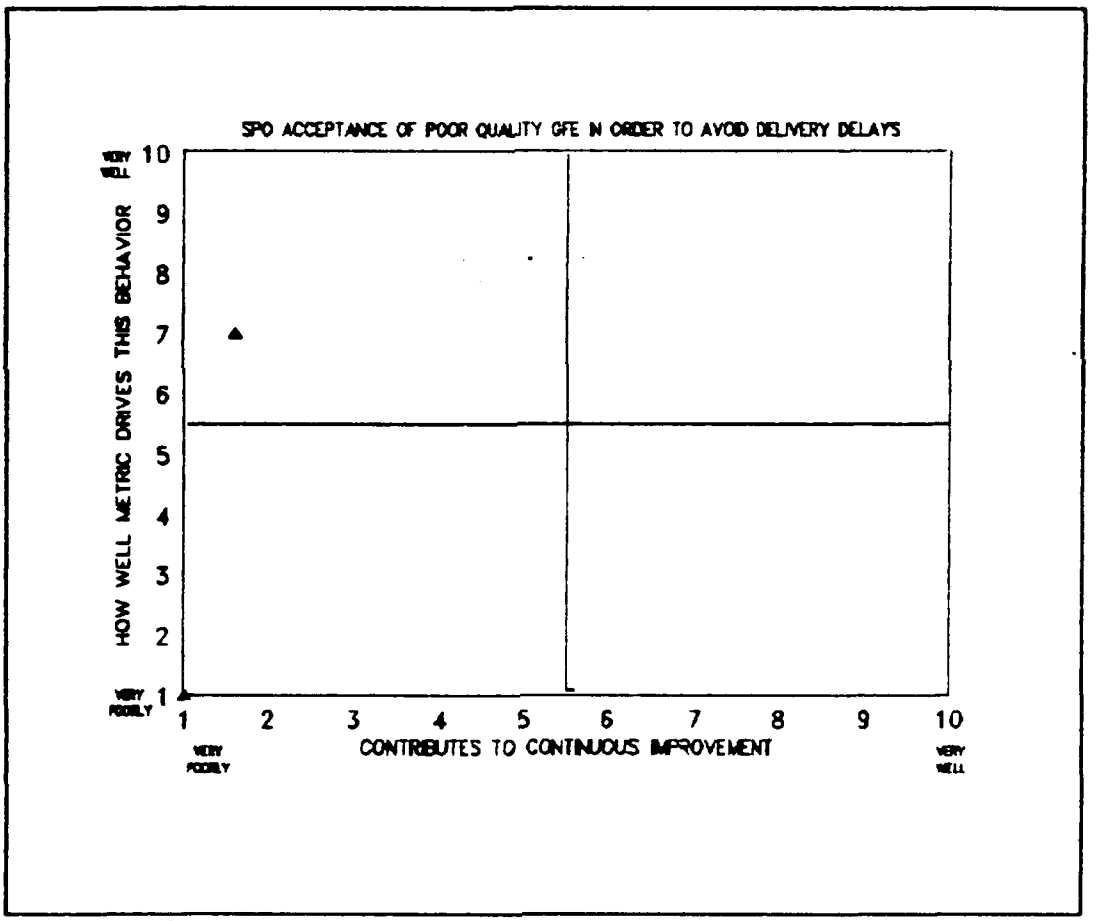


Figure A-97. Metric Seven, Behavior Seven

Behavior 8. Government will work GFE deals off-line until it's sure we can deliver

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	-	3	1	-	7.60
How well does behavior contribute to CI?	-	-	-	1	3	-	1	-	-	-	5.20

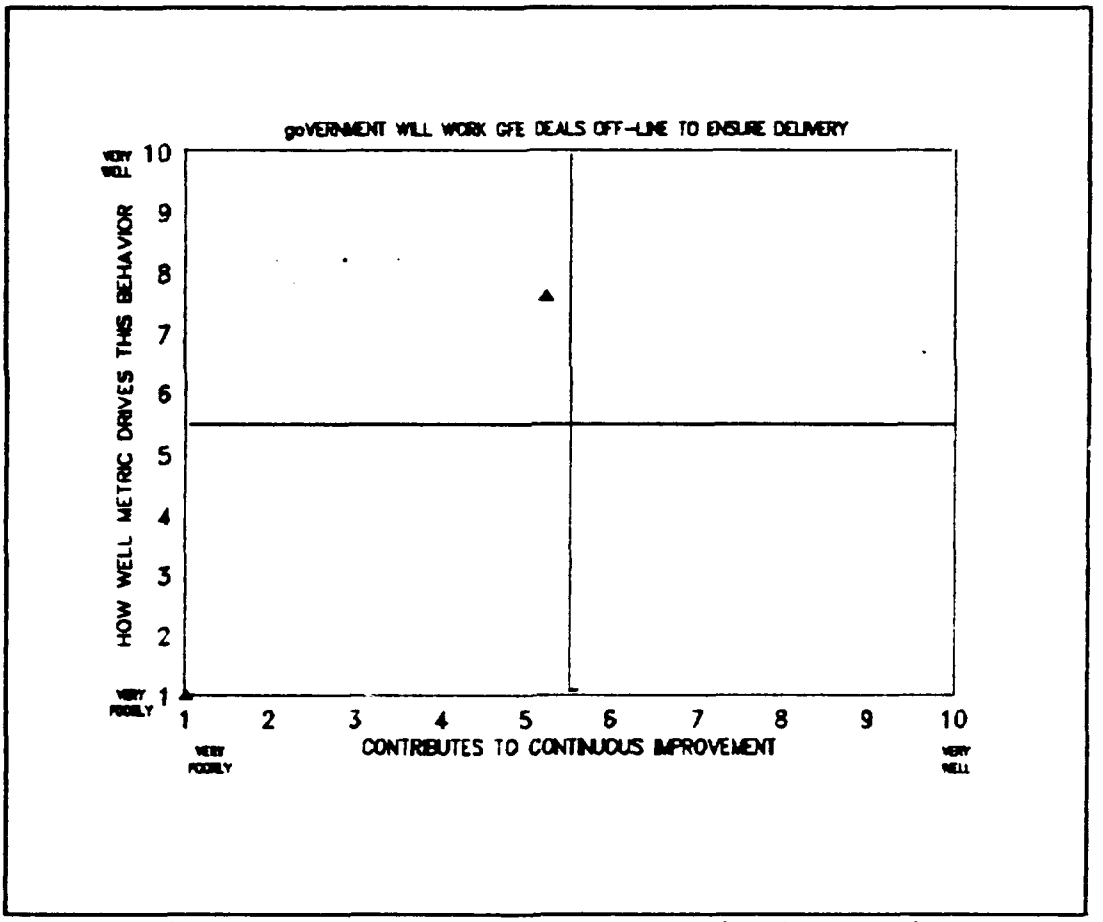


Figure A-98. Metric Seven, Behavior Eight

Behavior 9. Increased attention to ensure GFE producers deliver on-time

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	-	1	-	1	2	1	8.40
How well does behavior contribute to CI?	-	-	-	-	1	1	1	-	-	2	7.60

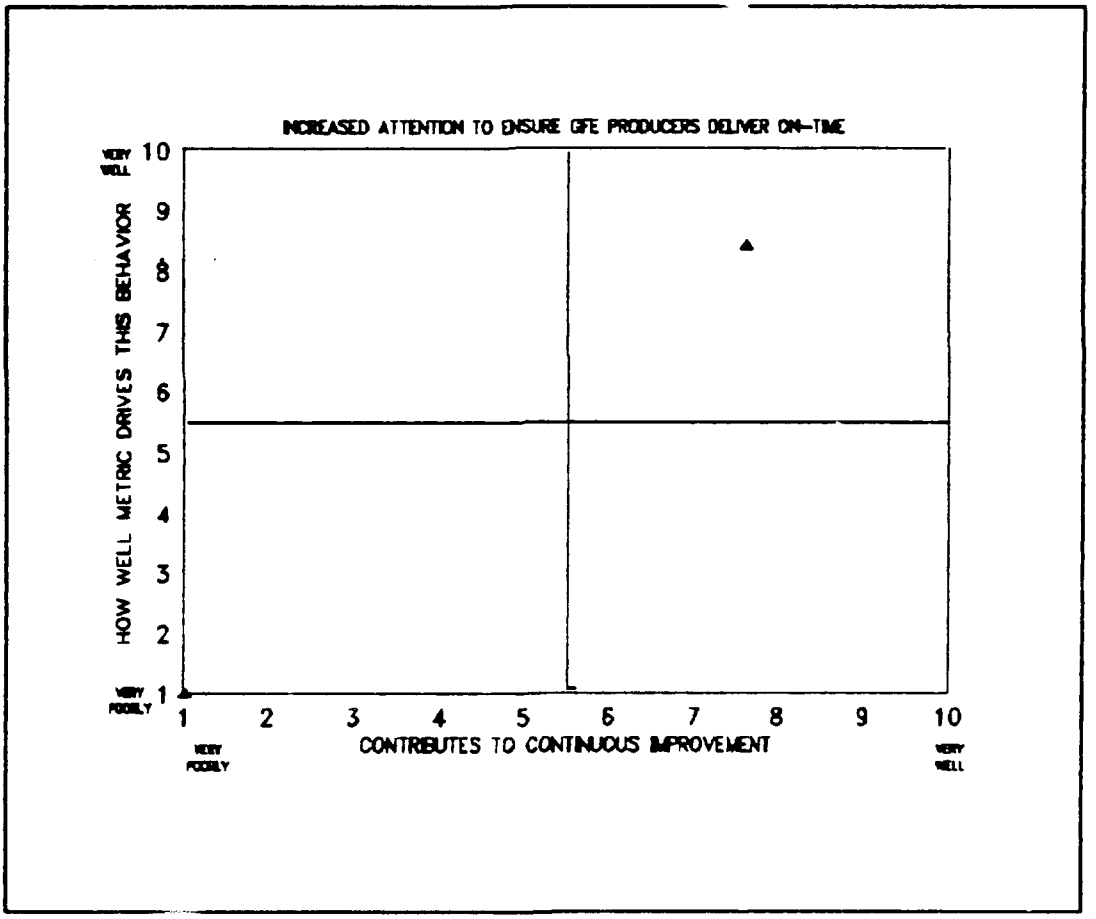


Figure A-99. Metric Seven, Behavior Nine

Behavior 10. The SPO will develop an efficient GFE production/delivery process

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	1	-	-	-	2	1	1	7.80
How well does behavior drive CI?	-	-	-	-	-	-	-	-	2	3	9.60

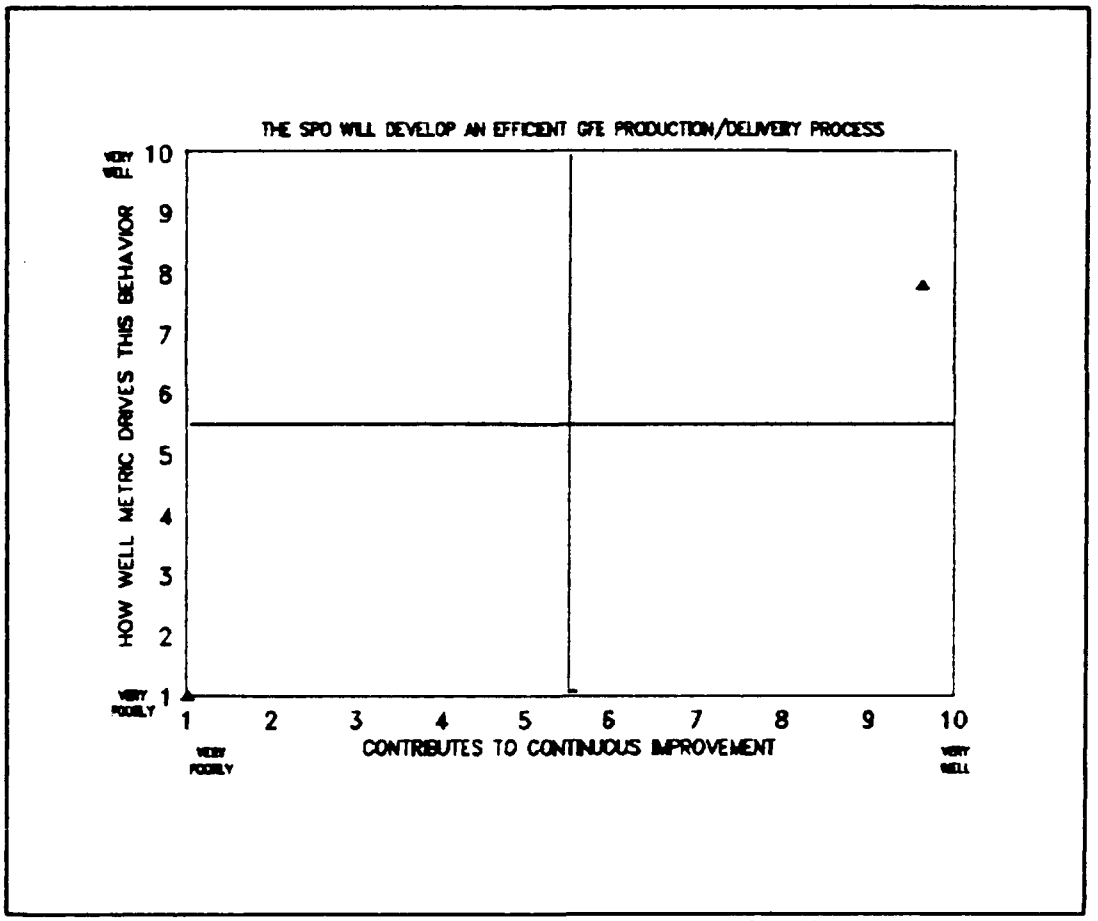


Figure A-100. Metric Seven, Behavior Ten

Behavior 11. Contractors will increase incoming GFE inspections

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	1	-	2	2	-	-	7.00
How well does behavior contribute to CI?	-	1	-	1	-	1	2	-	-	-	5.20

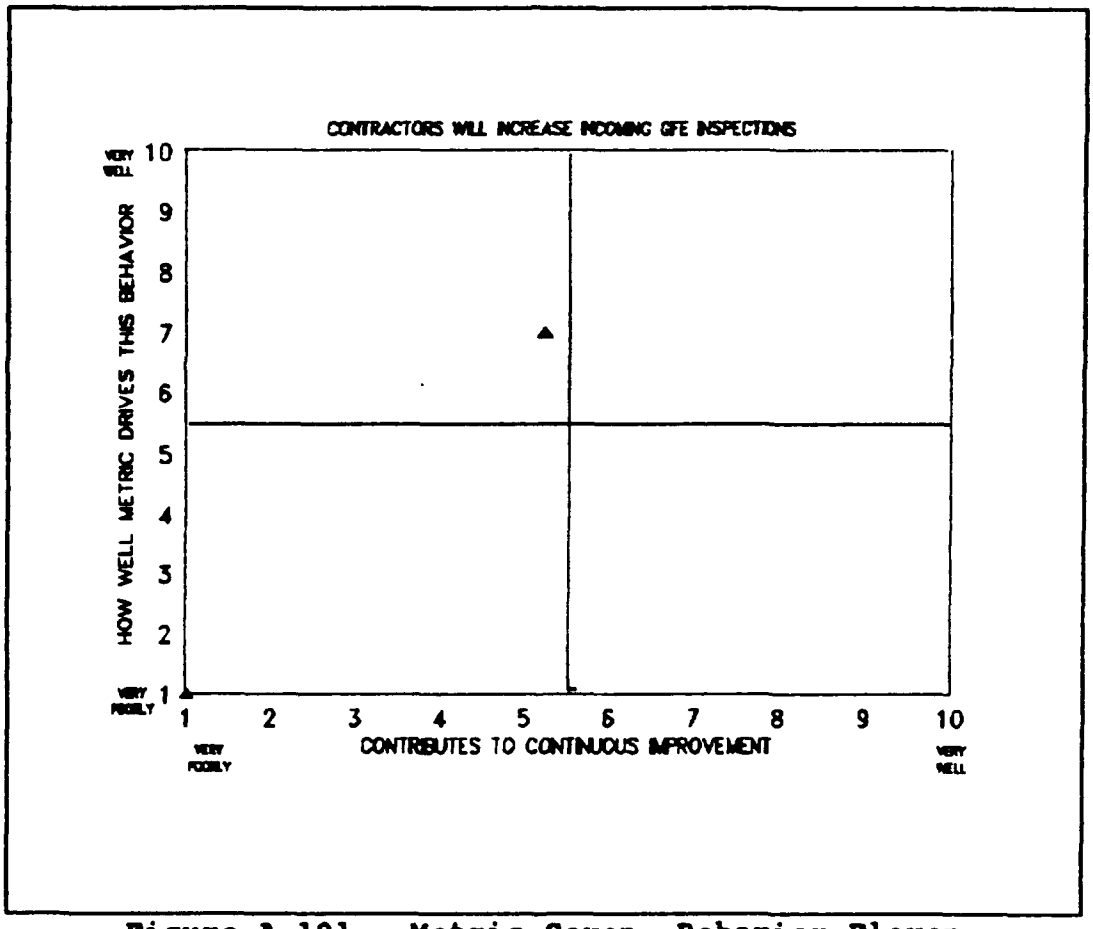


Figure A-101. Metric Seven, Behavior Eleven

Behavior 12. SPO tendency to reduce the amount of GFE on a project

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	1	-	-	-	1	2	1	-	7.00
How well does behavior contribute to CI?	-	-	-	1	3	-	-	1	-	-	5.40

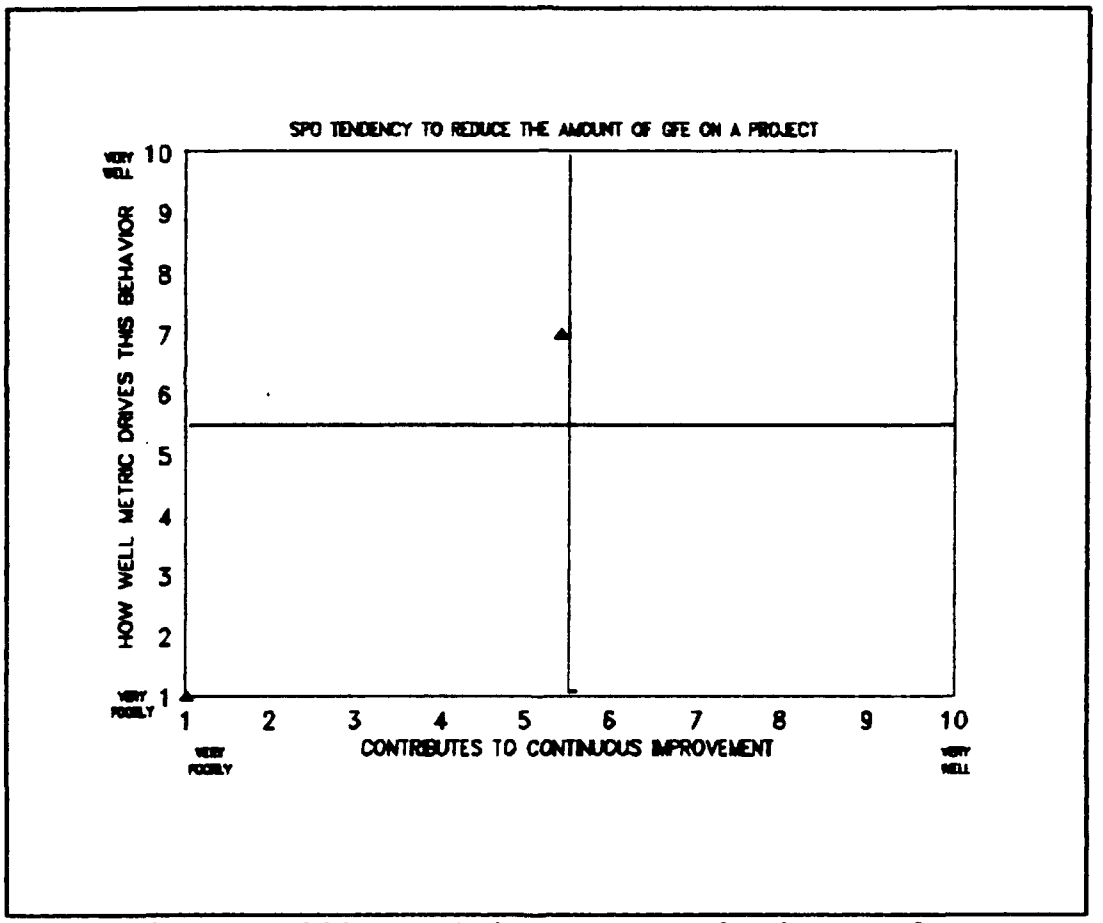


Figure A-102. Metric Seven, Behavior Twelve

Behavior 13. We will use quasi-black-market tactics to get GFE

** Criteria **	Participant Ratings 1	2	3	4	5	6	7	8	9	10	MN
How well does the metric drive the behavior?	-	-	-	-	2	1	-	1	-	1	6.80
How well does behavior contribute to CI?	1	1	1	-	1	-	-	1	-	-	3.80

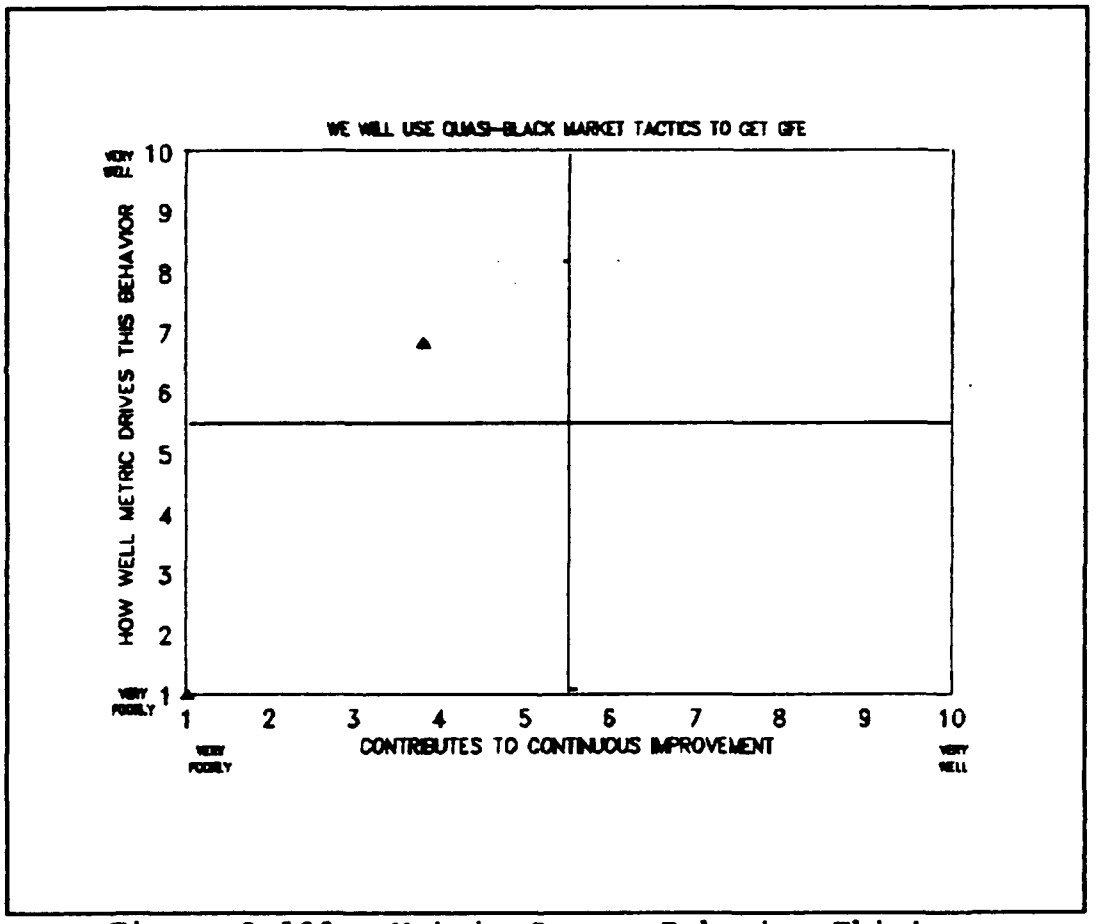


Figure A-103. Metric Seven, Behavior Thirteen

Appendix B. Evaluation Session Help Aids

List of Metrics.

1. Number of units of product actually delivered vs. the number of product required to be delivered on a monthly basis.

Delivery in this case is to the user. Required means contractually required. Units of product can be aircraft, pods, radios, or any other hardware or software item. This metric results in two numbers used for comparison.

2. Average number of days to respond to and resolve customer service report requests on a monthly basis.

Service reports are problems/issues identified for action by the users. Response in this case would be contractor/government resolution of the problem. An average is used because there may be numerous reports generated per time period.

3. Number of undefinitized contractual actions (UCAs) definitized within 180 days vs. those definitized after 180 days.

UCAs are actions which do not have agreed on price. Work is allowed to start using a not-to-exceed price so as to limit government cost risk. Used because it is often desirable to begin the work immediately rather than wait until a formal written agreement is reached. A UCA becomes definitized when subsequent agreement is reached on the exact changes to contract pricing.

180 days is a requirement mandated by federal law.

This results in two numbers: number definitized within 180 days and the number definitized after 180 days.

4. Number of project baseline breaches per month.

Breach means that the schedule in the finalized baseline was not met. In this case, baseline refers to an informal agreed upon plan of attack to meet cost, schedule, and performance goals (as opposed to the formal baseline for which breaches must be reported to higher headquarters).

Results in single number.

5. Average variance in days between the scheduled release of Request For Proposals (RFP) and the actual RFP releases by month.

Variance = number of days between planned release to prospective contractors and the actual release to prospective contractors.

An average is used because there may be multiple RFPs released per time period.

6. Number of contractor data submittals due versus those actually received on a monthly basis.

Contractor data submittals can be hardware/software specs, test reports, operation manuals, cost data, etc.

Received refers to receipt by Configuration Management within the SPO. Due means contractually specified and the actual date due.

Results in two numbers for comparison.

7. Number of actual Government Furnished Equipment (GFE) deliveries provided vs. the number actually required on a monthly basis.

GFE = owned by government but required under contract to be furnished to the contractor for using during the acquisition process.

Results in two numbers for comparison.

Software Help Summary.

1. In Topic Commenter mode you will see a stack of index cards containing text (metrics, questions, etc.). Selection of a specific card is accomplished by using the up and down arrow keys and pressing ENTER. The highlighted choice will show black text on a gray background.

Helpful Keys:

F1 - Help

Shift F1 - view the Participant Instructions as well as the question being considered.

F8 - view the text of the metric. ESC takes you back to the stack of index cards.

ENTER - selects a metric. At this point comments can be entered concerning the metric and the question(s) being asked.

F9 - after typing a comment on a card, this comment can be sent to all other participants by using this key.

F8 - allows you to view all participant's comments that have been sent as well as your own.

ESC - takes you back to the stack of comment cards.

Alt F9 - exits Topic Commenter mode. Do not use this until instructed.

2. In Alternative Evaluator mode, you will be rating behaviors against two criteria (on a scale of 1 to 10).

Helpful Keys:

F1 - Help

SHIFT F1 - view instructions along with the metric being evaluated.

ENTER - brings up rating scale. At this point you will have a behavior at the top of the screen and a criterion just below it. This is what you will use to perform the rating. To select a number simply use the left and right arrow keys. ENTER records your choice. You will see the second criterion displayed under the same behavior. You will perform the rating for this criterion in the same manner.

After rating a behavior against both criteria, the next behavior to be rated will automatically be displayed at the top of the screen.

F8 - view the original metric.

F7 - shows rating summary (how many behaviors and criteria for each that have been rated up to that point. The goal is to rate 2 criteria for each behavior.)

F3 - sends your final ratings to be compiled with the other participant's ratings. Use this key only after you have completed rating all the behaviors for a particular metric.

Alt F9 - exits Alternative Evaluator mode. Do not use this until instructed.

Note Taking Device.

1. WHAT ARE THE LIKELY BEHAVIORS THAT RESULT FROM THIS METRIC?

2. HOW WELL DOES THIS METRIC DRIVE THIS BEHAVIOR?

3. HOW WELL DOES THIS BEHAVIOR CONTRIBUTE TO CONTINUOUS IMPROVEMENT?

Bibliography

1. Aeronautical Systems Division. ASD Metrics Handbook. Wright-Patterson AFB OH: HQ ASD, 30 April 1992.

2. Aeronautical Systems Division. ASD Objectives and Metrics. Wright-Patterson AFB OH: HQ ASD, 1991.

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

4. Air Force Systems Command. The Metrics Handbook. Andrews AFB MD: HQ AFSC, August 1991.

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

6. Beery, Capt Warren, APTS Coordinator. Personal interview. HQ ASC, Wright-Patterson AFB OH, 27 Feb 1992.

7. Caudle, Capt Mark D. An Analysis of Total Quality Management in Aeronautical Systems Division. MS Thesis, AFIT/GSM/LSG/91S-6. School of Systems and Logistics, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, September 1991 (AD-A246661).

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

9. Cumberland Group and TQ Office, ASD. Measurement - The Handle On Your Business! March 15, 1991.

10. Dabrowski, Maj Ronald H., USAF. International Competitiveness - National Security Perspective. Maxwell AFB AL: Air University Press, 1990.

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

12. Department of Defense. Total Quality Management Guide for Implementation. DOD 5000.51G. Office of the Deputy Assistant Secretary of Defense for TQM, OASD(P&L) TQM, Washington DC, February 1989.

13. Downey, Tim. "DOD Accelerating into Total Quality Turn," Skywriter, p. 14, 15 February 1991.

14. Drucker, Peter. "The Deadly Sins in Public Administration," Public Administration Review, 40: 103-106 (Mar/Apr 1980).

15. Emmelhainz, Lt Col Larry W. "TQM Principles and Measures: Key to Successful Implementation," Air Force Journal of Logistics, 15: 34-37 (Summer 1991).

16. Emory, C. William and Donald R. Cooper. Business Research Methods. Homewood IL: Richard D. Irwin, Inc., 1991.

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

18. Hare, Paul A. Creativity in Small Groups. Beverly Hills CA: Sage Publications, 1982.

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

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

21. Juran, J. M. Juran on Leadership for Quality. New York: The Free Press, 1989.

22. -----. Juran on Planning for Quality. New York: The Free Press, 1988.

23. Kosinski, Capt Randy, Special Assistant to the Commander. Personal interview. HQ ASC/CY, Wright-Patterson AFB OH, 1 May 1992.

24. Linhart, Capt Chris, TQ Training Administrator. Personal interview. HQ ASC/CY, Wright-Patterson AFB OH, 4 June 1992.

25. Lobbestael, Colonel Wayne and Major Bud Vazquez. "Measure to Improve - Ideas on Implementing Measurement," Program Manager, 20: 29-34 (May-June 1991).

26. Main, Jeremy. "How to Win the Baldridge Award," Fortune, 121: 101-116 (April 23, 1990).

27. Meister, David. Behavioral Analysis and Measurement Methods. New York: John Wiley and Sons, 1985.

28. Milakovich, Michael E. "Total Quality Management In the Public Sector," National Productivity Review, 10: 195-213 (Spring 1991).

29. Oakland, John S. Total Quality Management. New York: Nichols Publishing Company, 1989.

30. Public Affairs Office, Wright-Patterson AFB OH. Legacy, Marcoa Publishing, San Diego CA: 1990.

31. Segalla, Ellen. "All for Quality, Quality for All," Training and Development Journal, 43: 36-45 (September 1989).

32. Ventana Corporation. GroupSystems. Dealer brochure. Tucson AZ, undated.

33. Virginia Productivity Center. Managing Quality and Productivity in Aerospace and Defense. Contract No. MDA903-85-C-0237 Managed by Defense Systems Management College, Fort Belvoir VA, November 1989.

34. Weaver, Charles N. TQM: A Step-By-Step Guide to Implementation. Milwaukee WI: ASOQ Quality Press, 1991.

Vita

Capt Robert J. Hayes was born on 26 September 1961 in El Paso, Texas. He graduated from Sycamore High School in Cincinnati, Ohio in 1979 and attended Miami University in Oxford, Ohio, graduating with a Bachelor of Science in Aeronautics and a Bachelor of Arts in Aeronautics-Mathematics. He attended Officers Training School in October 1984 and was commissioned 24 January 1985. After his commissioning, he served his first tour of duty at Nellis AFB, Nevada in the field of logistics. At Nellis, he served as the installation War Reserve Materiel (WRM) Officer and the assistant base mobility officer. There, he was responsible for the largest WRM and mobility programs in the Tactical Air Command. In May 1989, he cross-trained into the systems acquisition career field and was assigned to Wright-Patterson AFB, Ohio. He received his regular commission in August 1989. At Wright-Patterson, he served as program manager for the Tacit Rainbow Rotary Launcher in the Joint Tactical Autonomous Weapons System Program Office. He entered the School of Systems and Logistics, Air Force Institute of Technology, in May 1991.

Permanent Address: 8017 St. Matthew Dr.
West Chester, Ohio 45069

Vita

Captain Lawrence M. Miller was born on 18 April 1958 in Bad Axe, Michigan. He graduated from Farragut High School in Knoxville, Tennessee in 1976 and then attended the University of Tennessee at Knoxville, where he majored in Accounting. After graduating with honors with a degree in Business Administration in June 1981, he was employed in private industry. In July 1986 he entered Officer Training School, and was commissioned on 2 Oct 1986. His first assignment was as Chief, Combat Crew Communications for the 2034th Communications Squadron, Mather Air Force Base, California. In 1988, he was subsequently assigned as Chief of Operations, 2034th Communications Squadron and served in this capacity until his selection to attend the Air Force Institute of Technology (AFIT). He entered the School of Systems and Logistics at AFIT in May 1991. Captain Miller and his wife, Margie, have four children: Trevor, age 10; Callie, age 8; Dane, age 5; and Trent, age 2.

Permanent Address: 640 Gulfwood Rd.
Knoxville, Tennessee 37923

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 1992	Master's Thesis

4. TITLE AND SUBTITLE
AN EVALUATION OF SCHEDULE METRICS USED WITHIN AERONAUTICAL SYSTEMS CENTER

5. FUNDING NUMBERS

6. AUTHOR(S)
Robert J. Hayes, Capt, USAF
Lawrence M. Miller, Capt, USAF

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

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/GSM/LSY/92S-12

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

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

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This study focused on a selected group of schedule metrics in use at Aeronautical Systems Center (ASC) acquisition program offices. Over 300 metrics were collected from the System Program Offices (SPOs). The metrics data was sorted into the categories of cost, schedule, and performance. In order to narrow the scope of the project, the team decided to focus on schedule metrics. Seven of the most common schedule metrics were selected for evaluation by a group of five experienced acquisition professionals. The group was asked to input the likely behaviors driven by the metrics being addressed by the study. Next, they were asked to rate the metrics through a Group Support System at Armstrong Laboratory, Human Resources Division. The evaluation group rated how well the metric-driven behaviors contributed to continuous improvement. Results of the study showed a wide spread of behaviors, both positive and negative, that would likely be driven by the metrics. Most of the metrics rated were found to need improvement in terms of influencing behaviors that would lead to continuous improvement. However, many good metric-driven behaviors were identified and could prove helpful to program offices undergoing the challenge of developing their own internal metrics.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Acquisition, Procurement, Measurement, Systems Management	220
	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/LSY/92S-12

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: AFIT/LSC, Wright-Patterson AFB OH 45433-9905.

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. Slighly Significant d. Of No Significance

5. Comments

_____ _____

Name and Grade Organization

_____ _____

Position or Title Address

DEPARTMENT OF THE AIR FORCE
AFIT/LSC Bldg 642
2950 P St
45433-7765

OFFICIAL BUSINESS

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 1006 DAYTON OH

POSTAGE WILL BE PAID BY U.S. ADDRESSEE

Wright-Patterson Air Force Base

AFIT/LSC Bldg 642
2950 P St
Wright Patterson AFB OH 45433-9905