Measurement and the eSourcing Capability Model for Service Providers v2

21 January 2005
CMU-ISRI-04-128
Pittsburgh, PA

Mark C. Paulk, *Carnegie Mellon University*
Shari L. Dove, *Accenture*
Subrata Guha, *Satyam Computer Services Ltd.*
Elaine B. Hyder, *Carnegie Mellon University*
Majid Iqbal, *Carnegie Mellon University*
Kathleen O. Jacoby, *Accenture*
David M. Northcutt, *IBM Global Services*
George E. Stark, *IBM Global Services*

IT Services Qualification Center
Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3891 USA
itsqc.cs.cmu.edu

Abstract
The eSourcing Capability Model for Service Providers (eSCM-SP) is a "best practices" capability model. It provides guidance to service providers on improving their capability across the sourcing life-cycle, and it provides clients with an objective means of evaluating the capability of service providers. Measurement is fundamental to effective service management, business process outsourcing (BPO), and organizational improvement. This report describes the importance of aligning an organization's measurements with its business objectives. It provides guidance on measurement principles and defines four measurement categories that span the Practices of the eSCM-SP. Examples and lessons learned are used to illustrate the principles and guidance.

Keywords
eSCM-SP, eSourcing Capability Model, service provider model, quality models and systems, capability models, business process outsourcing, IT-enabled outsourcing services, IT-enabled services, outsourcing, outsourcing models, sourcing, measurement, return on investment, ROI, cost, effort, nonconformance, status, progress, performance, value-added.

Contributors
Stephanie Hamill, Mellon Financial Services; Chet Mauk, IBM Global Services; and Alan Yamamoto, IBM Global Services reviewed and contributed valuable insights during the development of this report. The overall document design was done by Paul Burke and Ken Mohnkern; the graphics were designed by Paul Burke. Editorial and technical writing support was provided by Ken Mohnkern.

Acknowledgements
This effort was made possible through the assistance of many individuals. First and foremost, the ITsqc Advisory Board and its Chair, Raj Reddy, provide ongoing guidance: Paul Byrnes, Integrated System Diagnostics; Betty Deimel, Gateway Associates; Shari Dove, Accenture; Tony Macina, IBM Global Services; Robert Parkinson, Mellon Financial Services; Ron Radice, Software Technology Transition; Prabhuu Sinha, Satyam Computers, Ltd.; and Shelly Williams, EDS. This work was enabled by our research and administrative staff, especially Ava Cruse.

Table of Contents

Preface **6**

1. Introduction **8**

2. The Service Provider's Context for Measurement **11**
2.1. Measurement Across the Sourcing Life-cycle 11
2.2. A Measurement Roadmap 12

3. Identify Objectives **14**

4. Define Measures **15**
4.1. Establish a Measurement Program 15
4.2. A Balanced and Comprehensive Set of Measures 18
4.3. Leading Versus Lagging Indicators 18
4.4. Operational Definitions 19
4.4.1. Service Delivery Unit 21
4.4.2. The Granularity of Measurement 22
4.5. Measurement Categories 22
4.5.1. Cost/Effort 23
4.5.2. Status/Progress 25
4.5.3. Nonconformance 26
4.5.4. Performance/Satisfaction 27

5. Implement Measures **29**
5.1. Collect and Process Data 29
5.2. Types of Measurement Analysis 30
5.3. Measurement at Capability Level 1 32
5.4. Measurement at Capability Level 2 32
5.4.1. Engagement Objectives – prf01 33
5.4.2. Verify Service Commitments – del04 34
5.4.3. Verify Processes - prf02 34
5.4.4. Resource Consumption – knw08 35
5.4.5. Summary of Level 2 Measurement 35
5.5. Measurement at Capability Level 3 36
5.5.1. Organizational Objectives – prf04 37
5.5.2. Review Organizational Performance – prf05 38
5.5.3. Process Assets – knw04 39
5.5.4. Summary of Level 3 Measurement 39
5.6. Measurement at Capability Level 4 40
5.6.1. Capability Baselines – prf08 41
5.6.2. Benchmark – prf09 41
5.6.3. Summary of Level 4 Measurement 43
5.7. Measurement at Capability Level 5 43
5.8. Measurement Activities by Phase and Level 43

6. Measure the Impact of the eSCM-SP **44**
6.1. Economic Profitability 44
6.2. Other Value Criteria 45
6.3. Other Types of Analyses 46
6.3.1. Payback 46
6.3.2. Internal Rate of Return (IRR) and Return on Investment (ROI) 46
6.3.3. Economic Value Added 47
6.3.4. Balanced Scorecard 47
6.4. Summary 48

7. Conclusions **49**

8. References **51**

Appendix A: An Overview of the eSCM-SP v2 **54**
A.1. Rationale Behind Development of the eSCM-SP 54
A.2. Structure of the eSCM-SP 54
A.2.1. Sourcing Life-cycle 55
A.2.2. Capability Areas 55
A.2.3. Capability Levels 57
A.3. Capability Determination Methods 59

Appendix B: Considerations in Building the Measurement Report **61**
B.1. Using the Goal-Question-Metric Paradigm to Define Measures 61
B.2. Role Perspectives 61
B.3. Top-Down Approach: Value Added Measures 62
B.4. Bottom-Up Approach: Capability Area Measures 62
B.5. The Balanced Scorecard 62
B.5.1. Scope of the Balanced Scorecard at the Different Levels of the eSCM-SP 64
B.5.2. The Relationship between eSCM-SP Practices, Measurement Categories, and Development of the Balanced Scorecard Perspectives 65
B.5.3. eSCM-SP Practices and the Balanced Scorecard Perspectives 66
B.6. Goal-Driven Measurement Applied 66

Appendix C: Operational Definition Template and Sample Measurements **68**

Appendix D: Examples **72**
D.1. Example #1: Call Center Measures, Service Level Indicators, and Operational Measures 72
D.2. Example #2: Mainframe Measures Using Relationships to Generate Measures 73
D.3. Example #3: Server Backups Dealing with Counts and Rates 74
D.4. Example #4: Customer Satisfaction Measures the Quality of the Measure 76
D.5. Example #5: Security—Dealing with Infrequent Events 78
D.6. Example #6: Planning—Dealing with Changes to the Agreement 80

Appendix E: Measurement from the Perspective of Capability Areas **82**

Appendix F: Examples of Measuring Value **92**
F.1. A Level 2 Example—Implementing the Threat Management Practices 92
F.2. A Level 3 Example—Creating a Measurement Repository 93
F.3. A Level 4 Example—Deciding Among Alternative Investments 95

Appendix G: Measurement in Related Standards and Models **98**
G.1. Measurement in ISO 9001 (Quality Management Systems—Requirements) 98
G.2. Measurement in COBIT 99
G.3. Measurement in BS 15000 and ITIL 100
G.4. Measurement in Software CMM v1.1 102
G.5. Measurement in CMMI v1.1 105
G.6. Measurement in COPC-2000 107

Appendix H: Terms and Definitions **109**

Preface

The primary audience for this report on measurement guidance for the eSourcing Capability Model for Service Providers (eSCM-SP) is professionals responsible for eSCM-SP implementations and anyone wishing to understand better how measurement can support the control and improvement of various Practices in the Model.

This report addresses issues related to setting measurable objectives and establishing a measurement program that supports the business objectives of service providers of IT-enabled services. It also discusses methods for creating operational definitions that support measurement related to Practices in the Model.

Section 1 of this report provides a general introduction to the eSCM-SP and measurement. It describes the importance of aligning an organization's improvement programs with its business objectives, discusses relevant measurement work (such as Balanced Scorecard, Six Sigma®, Goal/Question/Metric (GQM), and Practical Software Measurement (PSM)), and provides the motivation for measurement.

Section 2 provides a general overview of measurement principles that are important in the eSCM-SP context. The principles are derived from goal-driven measurement, which is founded on the precept that measures should provide insight into attaining business goals. As a corollary, the quality of the product or service depends on the quality of the process used in building the product or providing the service, therefore good process management includes measurement to evaluate how well the processes are working.

Section 3 describes lessons learned in establishing a successful measurement program.

Section 4 contains measurement guidance for providers of IT-enabled services. This is the specific guidance that an organization adopting the eSCM-SP is likely to find of value. This report identifies and describes four measurement categories that form a basis for an organization to develop or revise a set of measures for use in managing and improving sourcing service provision. These measurement categories span the various aspects of business and process management that should be considered when selecting and defining industry and organizational specific measures.

Section 5 provides a discussion of the critical measurement Practices at each eSCM-SP v2 Capability Level.

Section 6 contains a discussion of analyzing the value-added by implementing Practices in the Model or the overall Model. An executive may wish to analyze value to determine whether a Practice or process (or the eSCM-SP itself) is adding more benefit than it costs to the organization.

Section 7 presents the conclusions of this report.

Appendix A provides a general overview of the eSourcing Capability Model for Service Providers (eSCM-SP) v2.

Appendix B describes the methodology used in developing these measures.

Appendix C contains examples of operational measures in various measurement categories.

Appendix D contains examples of how measures can be used by service providers.

Appendix E describes the importance of measurement categories relative to the various Capability Areas.

Appendix F provides examples of measuring value added at different Capability Areas.

Appendix G describes how measurement is addressed in a number of related models and standards: ISO 9001, COBIT®, BS 15000, ITIL℠, CMMI®, the Software CMM®, and COPC-2000®.

Appendix H describes the terms and definitions used in this report.

1. Introduction

For a sourcing agreement to be deemed successful, all parties to the agreement must accrue their expected benefits—business leverage for the client, along with economic and relationship success for the provider. By defining and communicating measurable objectives, the service provider can track performance against those objectives and manage the outcomes, thereby providing both the service provider and the client with the business-value benefits they expect. The objectives and performance targets for an engagement must be aligned with the service provider's vision, mission, and strategic plans, as well as the client's and any key supplier's objectives to increase the chance of success [Chang 1997] .

One of the greatest challenges facing sourcing leaders is to coordinate and align individual work group processes to pursue improvements in productivity and quality for the common good. Businesses have made ongoing efforts to systematically improve since the birth of modern industrial economy at the beginning of the twentieth century. From Taylor's work on scientific management to Shewart's statistical process control, and more recently to the work of quality experts such as Deming, Juran, and Crosby, our understanding of how people, processes, and technology interact to affect quality, customer satisfaction, productivity, and efficiency in doing work has steadily evolved [March 1996]. This understanding of the importance of process, quality, and "best practices" has broadened beyond the initial manufacturing and assembly line environment, to system design and development, and now the IT-enabled service industries.

This report is based on the eSourcing Capability Model for Service Providers (eSCM-SP) v2 [Hyder 2004, Hyder 2004a]. The eSCM–SP v2 was developed by a consortium led by Carnegie Mellon University's Information Technology Services Qualification Center (ITsqc). It has three purposes: (1) to give service providers guidance that will help them improve their capability across the sourcing life-cycle, (2) to provide clients with an objective means of evaluating the capability of service providers, and (3) to offer service providers a standard to use when differentiating themselves from competitors. Released in April 2004, the eSCM–SP v2 is composed of 84 Practices, which can be thought of as the "best practices" that are associated with successful sourcing relationships. Each Practice is assigned values along three dimensions: Sourcing Life-cycle, Capability Area, and Capability Level. The first dimension, Sourcing Life-cycle, is divided into Ongoing, Initiation, Delivery, and Completion. Ongoing Practices span the entire Sourcing Life-cycle, while Initiation, Delivery, and Completion occur in specific phases of that Life-cycle. The second dimension of the eSCM–SP, Capability Areas, provides logical groupings of Practices to help users better remember and intellectually manage the content of the Model. These ten groupings allow service providers to build or demonstrate capabilities in each critical sourcing function. The third dimension in the eSCM–SP is Capability Levels. The five Capability Levels of the eSCM–SP describe an improvement path that clients should expect service providers to travel. An overview of the Model is provided in Appendix A.

The eSCM-SP is a recent entry in a long line of models and standards aimed at improving the capability of organizations in developing products and providing services [Paulk

forthcoming]. Information technology (IT) has been crucial in transforming industrial organizations into global networked enterprises that depend on computers, software, and related services to be competitive. Many business models, services, and solutions rely on IT to be both feasible and profitable. The benefits of IT have led some organizations to invest heavily in their in-house IT capabilities, but many companies have found it advantageous to outsource selected business processes and support capabilities so they can focus on their core competencies [Prahalad 1990].

Successful sourcing relationships depend on diligent management by the client and by the service provider. The eSCM-SP, which is specifically targeted at service providers of IT-enabled services, combines Carnegie Mellon University's expertise in building process and quality models with industry partners' expertise in sourcing to identify practices that increase the likelihood of successful sourcing relationships.

Measurement provides a firm foundation for effective management that addresses five needs. First, measurement is necessary for defining and tracking service levels, which provide the objective criteria for establishing customer-supplier agreements. Second, identifying organizational performance trends, which enable proactive management, depends on measurement and analysis. Third, measurement supports the effective and efficient allocation of resources. Fourth, continual process improvement is best institutionalized based on measurable improvement, which is derived from identifying opportunities for improvement and the value gained from improvement. Fifth, industry studies based on valid data provide a foundation for making informed trade-offs in supplier selection and monitoring, establishing service level agreements, and doing risk management.

Organizations establishing improvement programs will find guidance on what categories of measurement should be considered in the sourcing context and how to implement an effective measurement program. Examples of selecting and implementing specific measures are presented. This report refers to service-specific standards for further details on specific measures; for example, call centers may use the measures described in COPC. It is not intended to replicate that level of detail. This report guides service providers in defining the key indicators to be tracked and managed for the purpose of achieving desired benefits from the eSCM-SP.

Measurement is important in all three dimensions of the eSCM-SP: Sourcing Life-cycle, Capability Areas, and Capability Levels. Organizations at Level 1 are expected to have a set of business objectives that set a context for engagements, although they may not address a balanced, comprehensive set of objectives. Level 2 organizations should be able to measure cost/effort, status/progress, nonconformances, and performance/satisfaction toward achieving engagement objectives that link client requirements to business objectives as appropriate. Level 3 organizations should be able to measure performance and trends in achieving organizational objectives that integrate business objectives, client requirements, and improvement objectives in a balanced and comprehensive manner. Organizations at Levels 4 and 5 should be able to predict and measurably improvement process capability in a statistically meaningful way.

Measurement guidance is also provided in this report to determine the value added for sourcing-related business processes, specific Practices or Capability Areas in the eSCM-SP, or the Model as a whole. This information will help organizations prove benefits realized via eSCM-SP implementation and aid in the decision making for prioritizing which aspects of the Model to adopt. It is assumed that best practices drive process improvements, which in turn drive improved results. Improvements should be visible in business results such as increased sales, more responsive delivery, or lower costs.

This report is based on the philosophy of goal-driven measurement (i.e., that there ought to be a direct logical link between the business objectives of the organization and the measures collected and analyzed as a basis for action). Best practices drive process improvements that lead to business results. Measurement drives behavior, therefore there should be a direct link between objectives and measures. Since measures drive behavior, they should reflect a balanced set of business objectives. From an eSCM-SP perspective, measures should be derived from engagement as well as organizational objectives. These, in turn, may be driven by client requirements (typically captured in service level agreements), business objectives (such as growth and profitability), improvement objectives, statutory and regulatory requirements, etc.

The eSCM-SP must be validated if it is to be widely adopted in the sourcing community. Model validation measures will ultimately show how adoption of the eSCM-SP has affected a change in the outsourcing industry: improvements in engagement success, relationship building and management, and predictable value delivery. Model validation is not addressed in this report, but is being addressed via a number of studies at Carnegie Mellon University. A series of reports on different aspects of validation is planned for future years.

2. The Service Provider's Context for Measurement

The cost/effort, status/progress, nonconformance, and performance/satisfaction measurements for each engagement are often aggregated at the organizational level to provide the organizational insight needed for action.

2.1. Measurement Across the Sourcing Life-cycle

IT-enabled sourcing relationships move through three main phases: initiation, delivery, and completion. Each phase has a set of milestones that generally mark an increase in the amount of engagement detail understood by the service provider. Figure 1 depicts the phase and milestone relationships for a typical sourcing relationship. The initiation phase commonly begins when a Request for Proposal (RFP) is received. The RFP usually proceeds through a "down select" milestone where the client chooses a subset of the bidding service providers to refine their bids with additional detail about the service required and allows them to perform due diligence. During due diligence the service provider teams analyze the client operational environment and business practices to gain a more detailed understanding of the scope and complexity of the service and current performance. Once the client has chosen a provider, a contract is signed.

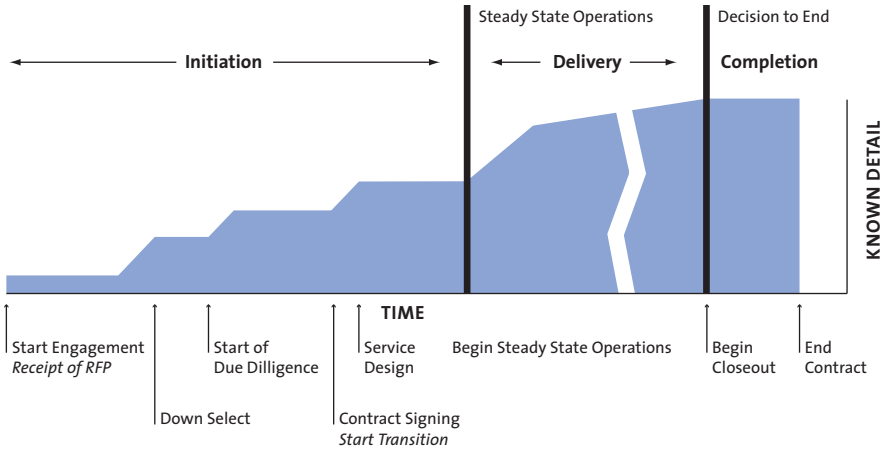

Figure 1
A View of the Sourcing Life-cycle

A transition period begins after contract signing. Service design may occur before or after signing the contract, depending on the kind of service desired. During the transition period, asset turnover and personnel transition take place. Typically, a new management team and structure are put in place. Such transitions are fraught with risks that must be managed, such as personnel turnover, loss of organization intellectual capital, and service performance degradation. The service provider's standardized processes and tools are implemented and client training occurs. As such, this is a period of intense learning—therefore learning curve effects have a major impact on both the client's and provider's organizational performance. The learning that takes place has to do with both the assimilation of an organization and its personnel into a new management structure along with the learning required by the organization to adopt and adapt to new processes.

After the transition period, the engagement moves into the steady state delivery period. This marks the end of the initiation phase and the beginning of the delivery phase. By being "beyond the learning curve" the service provider organization is readily able to achieve performance improvement targets that have specifically been set in the post-transition context. Additionally, as the relationship matures, performance targets set up in the initial agreement will typically no longer be valid because of either business change or technology change that was unforeseen in the initial agreement. Therefore, it is important that the service provider and client establish a link between business performance and technical performance on an on-going basis. This link leads to the development of new technical performance focus areas and targets.

As the contract comes to an end, the client may decide to end the engagement. The last phase of the Sourcing Life-cycle then begins. During Completion, the service provider and client work together to close out the contract and transfer previously agreed-to intellectual capital and assets (some of which may be measurement systems).

2.2. A Measurement Roadmap

Capability Levels in the eSCM-SP provide a set of plateaus that characterize improving capability, but it may take several iterations of improvement in order to achieve a Level. An organization may use the generic roadmap in Figure 2 to build its measurement capability for an engagement or organization. At all Levels, measurement should be driven by the organization's business objectives, and the data collected should support insight, control, or improvement.

The general template used for measurement in the eSCM-SP Practices includes identifying objectives, determining measures to track performance, collecting data on performance, periodically reviewing performance against the objectives, determining corrective actions to take when performance trends deviate from achieving the objectives, tracking the corrective actions, and tracking status/progress against the plan for the objectives.

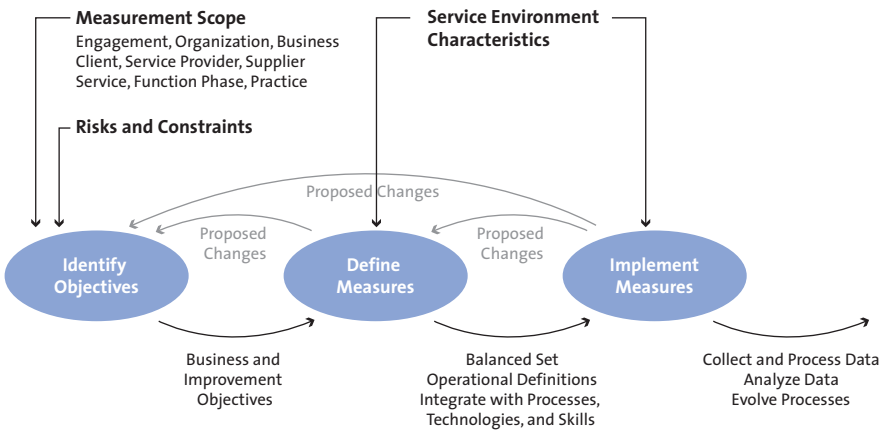

Figure 2
A Simple Measurement Roadmap

How do service providers, suppliers, and clients who want to benefit from measurement proceed? A number of organizations have defined fixed measurement sets with limited success. The key is to have a flexible approach that is adaptable for a variety of situations. Figure 2 illustrates an approach to establishing a service management measurement program based on the Issue-Measure-Indicator framework [McGarry 2002].

As shown in Figure 2, there are three major tasks with inputs from external sources. The first two activities, Identify Objectives and Define Measures, represent the planning part of establishing a measurement program. The third activity, Implement Measures, contains the execution part of the measurement program. There are feedback loops between each activity so that the program can remain a vital contributing process to meet the organization's needs.

The first task, Identify Objectives, is driven by the scope of the work effort and the associated risks and constraints. The scope of the work effort is bounded by at least four components:

- the organizational level (e.g., engagement or organization)
- the active participants (e.g., clients, service providers, and/or suppliers)
- the service being provided and the functions performed to provide the service
- the phase or eSCM-SP Practice being studied

Combinations of these attributes drive different issues and objectives to be measured. For example, at the engagement level, the objectives of the client and service provider may be different for an application management service from those of a help desk service.

The measurement scope boundaries are then combined with the specific risks and constraints that help identify and prioritize the objectives that the measurement system should help monitor. This prioritized list contains the business and improvement objectives that must have measures defined for the agreed-to scope.

The second task, Define Measures, involves selecting an economical and effective set of measures to address the key goals output from the Identify Objectives activity. These measures are examined to ensure that they are a balanced set for all stakeholders and goals. The operational definitions are established to eliminate ambiguity in the collection, analysis, and actionable nature of the measurement. Additionally, the approach to integrating the measures with the technology, skills, and procedures being used for the defined scope is implemented during this activity. Finally, the agreed-to measures are tested to ensure that they provide the visibility required.

The final task, Implement Measures, involves collecting, processing, and reporting the data for effective decision-making. During this activity it is common that additional questions are raised and new objectives or measures identified causing the process to iterate.

This roadmap is an abstract description of measurement that must be represented in different contexts. Each of these three tasks is elaborated in the following sections of this report.

3. Identify Objectives

Measurement is designed to provide data for decision making and those decisions and measures are driven by engagement and organizational objectives. Goal-driven measurement is the first step in the measurement roadmap [ISO 2002, Park 1996, McGarry 2002]. Organizations need to integrate objectives from many sources—business objectives, client needs, and improvement objectives—subject to constraints from the client, the business environment, and regulatory agencies. Engagement objectives are derived from client requirements and business objectives relevant to the engagement. Organizational objectives are derived from business objectives (e.g., increasing market share or revenue) and improvement objectives (e.g., decreasing response time or improving quality). Measurement is crucial for determining status/progress, characterizing effectiveness, and identifying opportunities for improved efficiencies.

The important thing for a service provider to do is to determine the set of measures to collect based on its engagement and organizational objectives, and then collect and use its measures consistently. A pre-defined set of measures from a model or standard is not required [Park 1996]. Here is a list of common areas where objectives are set:

- Finance/Budget focuses on financial management of the engagement and the ability to deliver the agreed to services within a budget.
- Customer Satisfaction/Loyalty focuses on critical attributes such as client retention, customer satisfaction ratings, and complaint resolution.
- Service Delivery Level Attainment focuses on the amount of work product, availability, or other service delivery unit delivered successfully in a given time period.
- Quality focuses on the objective and measurable aspects of the quality of services and products delivered.
- Time/Schedule focuses on critical service, product, or project time frames and the ability to deliver within the service commitment.
- Business Value focuses on the measurement of the attainment of the outcome of the sourcing agreement in terms of the client and service provider objectives.
- Human Resources focuses on changes to the skill inventory and internal job satisfaction.
- Productivity focuses on the efficiency of the production and delivery of the service.
- Process Conformance focuses on the execution of the agreed-to processes for delivering the service.
- IT Infrastructure focuses on the availability, threat, and event management of the IT infrastructure that supports the service delivery.

Each objective should be as measurable as possible, focusing on providing service to the client. The objectives relate to factors that influence client satisfaction or operational improvement, for example, "Increase client satisfaction by x%," and "Reduce process nonconformances by y%."

4. Define Measures

Establishing a set of common measures across an organization can simplify and leverage measurement deployment. An organizational measurement program is crucial for this deployment, and much has been learned on the factors affecting the success of such programs.

- The measures deployed across an organization should address a balanced and comprehensive set of issues important to business. An example of a strategy for defining such a set is in the Balanced Scorecard approach.
- A balanced set of measures should include leading indicators useful for control, as well as lagging indicators that support understanding outcomes.
- Repeatable, consistent measurement depends on good operational definitions. Without valid data, measurement can be effectively used for insight or decision making.
- There are four general categories of measures that are used: cost and effort, status and progress, nonconformance, and performance and satisfaction. Organizations typically collect data corresponding to each of these categories, plus others that may be of specific importance in their business environment.

4.1. Establish a Measurement Program

A good measurement program depends on performing the following:

- Have the appropriate sponsorship to define and deploy measures across the organization.
- Establish valid and consistent data collection mechanisms based on sound operational definitions.
- Identify key stakeholders who demonstrate commitment to establishing a balanced set of leading and lagging indicators that cover basic service categories, and use them to make timely decisions.

A measurement program requires executive sponsorship and commitment as well as resources to achieve the greatest business value possible. Measures should be linked to organizational and engagement objectives. Successful measurement cannot be a strategy unto itself. The fundamental precept of goal-driven measurement [Park 1996], the Goal-Question-Metric (GQM) paradigm [Basili 1996], and practical software measurement [McGarry 2002] is that measurement is driven by business objectives. This is critical to successful measurement.

Goal-driven measurement includes the following steps:

1. Identify the organization's business goals.
2. Identify the entities and attributes related to the goals.

3. Identify the questions and indicators that will provide insight into status, progress, effectiveness, or efficiency in achieving those goals.
4. Identify the data elements that should be collected to construct the indicators that will help answer questions.
5. Create clear, concise, and logical operational definitions of the measures.
6. Prepare a plan for implementing the measures.

Aligning measures with organizational objectives increases the likelihood of success greatly. Studies have shown that most measurement programs fail within two years. Researchers have identified a number of lessons learned that may be help service providers establish a measurement program [McGarry 2002, Gopal 2002, Kaydos 1999, Park 1996]. Eight important attributes of successful measurement initiatives are described below.

Start small

A service provider should start with measurement fundamentals and become more sophisticated as processes mature. It is best to start with a few measures that can easily and consistently be collected. Organizations are often overly aggressive in the beginning of a measurement program and overburden themselves early, when they would do better to collect a few key data points and then determine if and when more information is required. Too many organizations try to collect too many measures when their processes are not mature enough to support the data collection. This lesson is especially important for service providers with limited resources.

Provide adequate training

All those involved in the measurement efforts need to understand what a particular measurement and analysis represents so the data can be interpreted correctly and the appropriate decisions can be made. This implies that the training needs of the different people involved in defining measures, collecting and analyzing data, and making decisions will be appropriately considered. It also implies an understanding of variation since there is normal, random variation in every process.

Demonstrate commitment

Management commitment is necessary for any change to succeed in an organization, whether it is a measurement program, a process improvement initiative, or the diffusion of innovations. Commitment is demonstrated by providing adequate resources for measurement and using the measurement results in making decisions. The latter may be shown by supporting people whose reports, actions, and recommendations are backed up with data.

Minimize costs

A measurement program must be cost-effective in business terms. Automation should be considered where feasible, since the cost of manual data collection and analysis is likely to be high. Personnel may have a natural resistance to taking on data collection and analysis activities in addition to their regular duties.

Focus on action
In selecting which measures to collect, consider only those on which you can and will act. Ask yourself "What will I do if this goes up or down?" If the answer is "nothing," do not collect the measure. Not only is it a waste of time to collect measurements that will not be used for decision making, but there is a potential negative impact when the staff observe that measurement efforts are not providing value. In addition, if you do not act on information about problems or potential problems with your processes or products, you will lose credibility with stakeholders.

Focusing on action implies setting clear goals for the measurement program. Measurement is crucial to management, allowing decisions to be made based on facts, some of which may be unexpected. For example, it allows confirmation that the expected benefits of a process change are realized and that unintended side effects are avoided.

Do not use measurement to reward or punish the individuals collecting the data.
Measure the process and/or product, not the people. Assume that people are doing their best, and that the opportunities for improvement lie in the process, the training, etc. If measures are used to evaluate individuals, counterproductive behaviors are likely to arise. Deming strongly opposed performance appraisals because the system predetermines the bulk of performance [Deming 1986]. Austin has shown that measurement for motivational purposes will drive dysfunctional behavior in the absence of a comprehensive set of measures [Austin 1996]. Performance appraisals may be an unavoidable consequence of an organization's need to build the capability of its workforce, but integrating performance appraisals into the measurement program is a high-risk strategy.

Implement a balanced set of measures
Measures should reflect the engagement's and the organization's objectives and priorities since people will change behaviors when measures are implemented. Do not emphasize one measure or indicator to the exclusion of others. For example, if you measure productivity without tracking quality, the results will be fast production at the expense of quality. This is true for both process and product. When establishing the set of measures to be collected, ask yourself what behavior each measure will incite, since you get what you measure.

Communicate
Timely communication is essential. In some cases, "timeliness" impacts the day-to-day collection, analysis, and use of measurement data. In other cases, communication is crucial to building buy-in that measurement is an important, value-added activity within the organization. Feedback loops that articulate the impact of data on decisions clear to the people collecting and analyzing the data support the consistent and timely collection of valid data.

Once a measurement program is successfully established, it should have a number of positive attributes:

- Data collection is automatic and natural.
- Data are widely available.

- People seek data as a basis for decision making.
- Failure leads to understanding rather than blame.
- Numeric objectives are accompanied by rational plans.
- Improvements are made regularly to the measurement process.

4.2. A Balanced and Comprehensive Set of Measures

The Balanced Scorecard (BSC) defined by Kaplan and Norton [Kaplan 1996] is a tool that translates an organization's business objectives into a "balanced" set of performance measures in four key areas: Financial, Customer, Internal Business Processes, and Learning & Growth.

The BSC can be used to help an organization organize its measurement efforts and ensure that those efforts provide a balanced perspective. BSC measures do not need to be defined for every eSCM-SP Practice. Rather, an organization needs to determine the measures critical to successfully meet both its customer's and its own business goals and objectives, as well as measures to help manage the organization's business and eSCM-SP implementation. Further detail on using the BSC with the eSCM-SP is in Appendix B.

4.3. Leading Versus Lagging Indicators

Some measures reflect the current performance, failures, or exceptions related to a particular process or Practice by predicting the probability of success in meeting specific objectives. These measures serve as leading indicators that, based on rules or expert judgment, invoke pre-determined or ad hoc intervening action to favorably alter that probability. Leading indicators are likely to be used at the operations level of monitoring and control (e.g., service level agreement (SLA) measures). They are also likely to be included in exception reports and escalation notices.

The results or value produced by some processes or Practices cannot be ascertained, or do not make sense, until after a cycle is completed. Such instances may require measures that are lagging indicators of performance, and that will invoke intervening action between two cycles. These measures are typically used in periodic reviews of performance necessary for regular assessment and reporting, as may be required by clients or by regulatory authorities. These performance measures are therefore utilized more in tactical management and control.

Not all performance can be directly measured or quantified with data collected from the work products of processes or Practices. In such situations, qualitative measures may be required to provide management the feedback necessary for improvement and control. Even when quantitative measures are readily available, they may not accurately or completely reflect the value, or the lack of it, produced from the perspective of the beneficiaries. It is common, for example, to find service providers who have successfully met SLA targets, but who are bewildered by unhappy customers or service personnel, the two key constituencies within the scope of the eSCM-SP. In most of these instances

the quantitative performance measures have either been misleading or have provided too narrow a perspective to be useful. Customer satisfaction measures may therefore be required in addition to the usual set of performance measures. While satisfaction may be experienced and expressed in qualitative terms, a well-designed numerical index can help translate such feedback into quantitative information useful for analysis and communication. These indices are often required by contractual obligations and tied to financial terms and conditions. Even when they are not, such tools for measuring satisfaction are useful in validating the views presented by other types of measures. They are particularly relevant to IT-enabled services because of the relatively larger element of human activity and interaction in this category of services.

These indices, however, need to be periodically reviewed to ensure that they accurately model satisfaction where there have been noticeable, significant changes in customer or employee expectations, attitudes, or perceptions, as often occurs after organizational changes such as mergers, acquisitions, transition of employees (e.g., as part of sourcing deals), and changes in key personnel. Not doing so would introduce distortion between the levels of satisfaction actually expressed, and the levels perceived by service provider management.

4.4. Operational Definitions

The first step in establishing a quantitative understanding of a process is understanding how it is measured. Process consistency and data validity can then provide a basis for rigorous analysis. It is common for organizations to discover that measures are not as consistently defined, that data are not as consistently collected, and that the processes used are not as consistently implemented as desired. Well-defined processes and measures are a prerequisite to management by fact. Much measurement work has been aimed at building frameworks for establishing good operational definitions for such fundamental measures as effort [Goethert 1992].

Operational definitions are the rules and procedures used to capture and record data [Park 1996]. They determine what the reported values include and exclude. They should meet two criteria:

- Communication: will others know what has been measured and what has been included and excluded?
- Repeatability: would others be able to repeat the measurements and get the same results?

Some measures will be objective; some will be subjective and dependent on human judgment. While objective measures may be preferable in many contexts, subjective measures may be the best choice in areas such as client perception and customer satisfaction. Operational definitions should ensure consistency, repeatability, and the minimization of measurement errors and "noise" in the data.

Operational definitions are clear, concise, detailed explanations of measurements that are fundamental to data collection. It is easy to assume that the people collecting the

data understand how to complete the task. However, people have different views and opinions, and these will affect the data collection. For example, erroneous data collection occurs when deskside-support staff operate with a different view of what constitutes a closed ticket for each customer interaction. Similarly, when healthcare reports are being checked for errors, the data collection will be meaningless if the definition of "error" has not been specified. The only way to ensure consistent data collection is by means of detailed operational definitions that eliminate ambiguity in a particular environment. When collecting data, it is essential that everyone in the system has the same understanding and collects data in the same way. Operational definitions should therefore be specified before data collection begins.

An operational definition has four main components:

1. The characteristic of interest.
2. The measuring instrument. This is usually either a physical piece of measuring equipment such as an accounting system, performance monitor, or clock; or alternatively, a visual inspection.
3. The method of the test. This is the procedure for taking the measurement and making the calculation, if needed (e.g., sampling/collection procedure).
4. The decision criteria. This includes examples of acceptable versus unacceptable values.

The start and finish points for time-based measurements need to be clearly defined. Additionally, the units and degree of accuracy should be specified for all measurements.

Table 1 lists the steps to be taken when creating an operational definition, with examples.

Table 1
Steps in Creating an Operational Definition

Step	Activity	Example Measure	Example for Step
1	Identify the objective of the measure.		Explicitly state the overall objective of the measure, as well as a specific question that the measurement is intended to answer. The eSCM-SP team has defined four typical measurement categories to help managers focus their objectives: cost/effort, status/progress, nonconformance, and performance/satisfaction.
2	Identify the Characteristic of Interest.	Actual cost per unit of service delivered over time	Determine the number of servers managed per full-time equivalent staff. The following service-related activities are included in the cost: equipment installation and physical management, change management, capacity management, performance tuning, basic problem management, provisioning and production control, and physical security.
3	Select the measuring instrument.	Actual cost measured using accrued accounting system charges Unit of service delivered	Measure the effort based on the hours charged by the server team through the organization's activity-claiming system. The financial team sets the number of hours expended for a full-time equivalent staff at the beginning of each fiscal year. This number is divided by 12 to arrive at the monthly FTE factor. The number of servers managed will be collected through a monthly visual inspection that identifies all of an engagement's servers in use by type (Intel, Unix, other).
4	Identify the method of test.	Divide the total amount of service delivered by the total cost expended during the time period of interest	The formula for computing the cost per unit of service delivered is (number of servers managed)/(hours charged/FTE factor). The result should be reported to the nearest tenth of a server (i.e., one decimal point accuracy).
5	State the Decision Criteria.		The measurement should trend up over time. In other words, as the business increases its capability, the number of servers managed per FTE should increase. If the trend moves in the negative direction, corrective action should be taken.

Terminology is critical to good operational definitions. Activities that build measurement concepts in a goal-driven measurement include the following:

- Base measures are collected at various points in the process.
- Derived measures are calculated from base measures.
- Indicators compare measures to decision criteria.
- Questions are answered with respect to indicators.
- Goals are achieved by basing decisions on the answers to the questions as appropriate (management by fact).

4.4.1. Service Delivery Unit

Many measures may be derived from base measures, frequently by normalizing the data against the Service Delivery Unit (SDU). The concept and definition of an SDU is essential to measurement, analysis, and control within the context of capability development and improvement based on the eSCM-SP. An SDU is the specific, identifiable, and measurable unit of value delivered to the customer by the service provider as an outcome of service provision. The following are examples of SDUs:

- support call placed at a help desk (technical support services)
- outbound call made to current or prospective customers (telemarketing services)
- processing of an insurance claim (financial services)
- recruitment of an employee (HR services)
- minutes of voice data to be transcribed (transcription services)
- maintained function points or lines of code (application management services)
- period of network access (telecom services)
- gigabytes of storage maintained per period (storage services)

The actual definition of an SDU may vary across business segments and categories of services. The idea, however, is the same. The consumption of one SDU is expected to provide customers some finite accountable benefit for which they are willing to be charged. Similarly, the service provider is expected to expend a finite amount of capabilities and resources for the provision of one SDU. It is important for service providers to clearly define and communicate the idea of an SDU across units of its organization involved in delivery and supporting the services.

The definition is also a basis for discussions and negotiations with clients on issues such as service levels, demands, capacities, and costs of services to be delivered. Definitions of SDUs are necessary for normalization to facilitate comparisons across clients, engagements and contracts and for baselines and benchmarks. They also enable activities such as cost-benefit analysis, sensitivity analysis, planning and scheduling, and investment analysis.

4.4.2. The Granularity of Measurement

This report provides guidance on the measures that can be used to determine if the implementation of eSCM-SP Practices is effective, efficient, and adding value. Providing industry with the means (measures) to determine if the best practices in the eSCM-SP are working is helpful, but identifying three or four measures per Practice implies several hundred measures that an organization should consider in implementing the eSCM-SP—an overwhelming number.

Many of these measures are very similar. For example, measuring the effort to implement a Practice may be important for determining status, effectiveness, or efficiency. The general concept of effort measurement is relatively simple; the issue is one of granularity of the activities to be measured.

From a business value perspective, it may not be useful to measure the effort expended in implementing every Practice in the Model. In some instances, it may make more sense to measure effort for a group of related Practices (e.g., the policy and procedure Practices on risk management within and across engagements). The Capability Areas within the eSCM-SP provide one potential grouping for measurement. Business processes within the organization provide an arguably more useful grouping, which may or may not correspond fairly closely to Capability Areas. From a goal-driven measurement perspective, the measures defined, the data collected, and the analyses performed should be those that support achieving business objectives.

4.5. Measurement Categories

In the development of this report, much consideration was given to the level of specificity that would be most appropriate. The aim, and the associated trade-off, has been to define categories of measures that would be specific enough to be of immediate value to implementation teams, yet generic enough to apply across a broad spectrum of categories and segments of IT-enabled services. A certain level of consistency is also expected with the use of common measures and indicators, allowing for comparative analyses at later stages.

In addition to adopting the goal-driven measurement approach to outline the categories and types of measures, a set of guiding principles was also drafted. Based on those principles, measures should have the following properties:

- relevance within the scope of eSCM-SP Practices
  - for monitoring performance or progress in implementing eSCM-SP Practices
  - for controlling performance or progress toward Practice objectives as described in the Practice rationales
- feasibility from a collection and analysis perspective
- small amount of data collection required (thereby minimizing the cost and effort)
- simplified data collection
- reduced time to analyze data and make decisions

Following this approach, this report outlines a common set of measurement categories that are considered useful for providers of IT-enabled services. These are general categories under which most measures can be classified, based upon their primary purpose and the type of visibility and control they provide to managers. Each category provides a major perspective for improvement and control applied at different levels of management. Measures from these categories are found to be useful not just in one or two specific Practices, but across larger sections of the Model. The scope of a given measure could be a particular Practice, or a group of Practices within a Capability Area. Appendix E provides guidance on the applicability of each of these categories across the eSCM-SP Capability Areas. A measure applied to a Practice or Capability Area can provide control feedback to one or more other Practices. This report discusses four measurement categories:

- cost/effort
- status/progress
- nonconformance
- performance/satisfaction

This section provides a high-level discussion of the objectives of each category of measures, including the general approaches and techniques that could be useful to implement and manage the measures. It is critical to reiterate that these are not required measures, but are rather recommended categories of measures that each organization should consider when developing its own specific measurement program. Examples of defining measures are provided in Appendix C.

Note that measurement occurs in a context and understanding the drivers and context for the data are crucial for making well-informed decisions. Simply knowing the value of a measurement often does not meet the analyst or decision maker's real needs. For example, average speed to answer a call (a call center performance measure) requires analysis of call volumes, talk times, and staffing levels to truly understand performance.

4.5.1. Cost/Effort

Cost/effort measures focus on the actual cost or effort the activity requires, as well as its associated variances.

Costs represent a major and commonly-implemented category of measurement, particularly under the fact-based management philosophy. While the specific types and definitions of costs may vary across organizations, the objectives for estimating, monitoring, and controlling costs are largely the same. Some of the common objectives for measuring costs, addressed within eSCM-SP Practices, include planning and budgeting, pricing decisions, operational efficiency, accounting of resources, project management (e.g., earned value), cost-benefit analyses, and identification of value enhancement. Other objectives for measuring costs include contractual obligations for reporting or transparency, performance evaluations (e.g., scorecards), and quality management initiatives (e.g., cost of quality). The effectiveness of cost-based measures may depend on the following:

- defining and understanding the cost categories and cost units
- understanding the effects of variances in costs on outcomes and objectives
- mapping costs to cost drivers (e.g., activities, processes, services)
- integrating performance evaluations and incentives
- feasibility and efficiency of accounting and reporting

The level of detail at which costs may be tracked or measured will depend on the organization's control and reporting objectives, and may be constrained by its complexity and span. In certain instances costs may serve as direct measures of progress, performance, and outcomes, while in other instances they may serve as indicators or surrogate measures, especially when variances in costs indicate problems or issues related to projects, processes, and activities. Since IT-enabled services, especially those that fall within the category of business process outsourcing (BPO), require a relatively high application of human knowledge and effort, a major cost driver is effort.

For the purpose of this discussion, effort includes all activity expended, directly or indirectly, toward fulfillment of service planning, delivery, and support objectives. Definition, identification, tracking, and measurement of effort is therefore valuable from a tactical and operations management perspective. While costs may serve as reliable measures or indicators of effort, the converse is often also true. Which primary view organizations adopt depends on their existing measurement and control set up, or their management approach or philosophy.

Effort-based measures can be useful to identify procedures, activities, or services that may be consuming a significantly higher amount of effort than estimated or required. In such cases, these measures serve as indicators of existing or potential problems in ongoing delivery and support of services. Further analysis can also lead to review of cost estimates and charges associated with processes or services that consume that effort. During initial stages, when services are being designed or planned for deployment, effort-based measures may be needed for analyzing staffing requirements (e.g., capacities by competencies), estimating duration of activities and tasks required for meeting specified (client) requirements, modeling costs to be recovered, and/or for identifying potential areas for automation.

Again, organizations may classify and prioritize types of effort based on the peculiarities of their business environment, the characteristics of services provided, their measurement and control objectives, and specific reporting requirements. Organizations also need to define and adopt a unit for effort (e.g., full-time equivalent) that is useful in planning, communication, and analysis. Fluctuations in currency exchange rates may also need to be considered for international work.

Finally, while cost/effort may be viewed as closely-related measurement inputs for managing service delivery and support, it may be useful to distinguish the roles they play in management. For example, tracking only effort may be ineffective if the cost per unit of

that effort is not predicted or tracked. A particular process may be consuming the expected number of person hours to complete, but if those hours are at a vastly different rate than expected, actual costs will not match plans.

Useful resources for building operational definitions of cost/effort measures include "Software Effort and Schedule Measurement" [Goethert 1992] and the Project Management Body of Knowledge [PMBOK 2000].

4.5.2. Status/Progress

Status/progress measures focus on the status and progress that the organization is making on meeting their plans (whether they are resource plans, program or project plans, or budgets).

While the steady-state delivery and support of services are ongoing streams of activities, events, and outcomes, other major aspects of providing services are planned and executed in the form of projects. These include work related to the design, development, deployment of service; changes implemented to improve service levels or eliminate problems; establishment of new or modified policies and procedures; conducting training; implementation and testing of controls; and compliance audits. Project-based activities related to this type of work need to be tracked for their status and progress to provide managers both a static and dynamic (progressive) view of their service organization. Status measures can be defined and organized to provide a snapshot of current activity or states. Progress measures can similarly be used to review the rate of change or degree of accomplishment since the last snapshot.

Both views are required to analyze effectiveness and efficiency in the implementation of plans and procedures, as well as to predict the probability of success in meeting objectives as required, or the associated risks of failure. The two views are also useful in periodic checks on conformance to policies and regulations. Intervening or control action may then be initiated, if necessary, to prevent poor performance, failure, or adverse consequences (e.g., penalties, fines, loss of customer value).

Measures for eSCM-SP Practices can be based on widely-used measures from the fields of project management and auditing. For example, measurements based on planned versus actual, or earned value (EV) (e.g., cost variance and schedule variance) can be applied to several Practices after clearly defining the meaning and implications of such variances within the scope of the Practice(s) or the Model. The degree or absolute value of the variances may have different meaning or impact across different domains of activity, services, or segments. The definition of the status or progress measure should therefore add value by helping managers interpret the condition or situation being reflected, and in determining the appropriate course of corrective action or reporting.

Planned versus actual measures typically support the goal of measuring service delivery processes. These measures typically give an understanding of how resource usage is tracking over time as services are delivered. Such measures are usually reported as the difference between the planned usage in a given time interval and the actual usage in that interval, with negative values usually indicating over-consumption and positive

values under-consumption. If absolute differences are used, it is important to include the actual planned level in any reports, because that provides the necessary context for the proper interpretation of the data. An alternative that avoids this problem is to report the percentage difference calculated as

$$\frac{\text{planned - actual}}{\text{planned}} \cdot 100$$

This approach normalizes the deviation and reduces the possibility of misinterpretation. Further, if planned versus actual measurement is to be tracked over time, normalization is necessary for any graphical display of the data.

Useful resources for building operational definitions of status/progress measures include "Software Effort and Schedule Measurement" [Goethert 1992] and the Project Management Body of Knowledge [PMBOK 2000].

4.5.3. Nonconformance

Nonconformance measures primarily address the goal of ensuring conformance with required policies, procedures, statutes, and regulations, while delivering and supporting services at the required service levels.

Differences in interpretation, judgment, or execution on the part of service personnel, with respect to service delivery and support activities, can introduce undesired but controllable variation in performances or outcomes. Common and consistent understanding and application of contracts, agreements, statutes and regulations are needed. Guidelines, policies and procedures are therefore required to reduce such variation and uncertainty, thereby avoiding unnecessary costs and risks.

Organizations need to identify, and reduce or prevent nonconformances or failures in meeting the requirements of policies, procedures, statutes, and regulations. This requires a set of measures and indicators that will provide managers the required visibility and control over processes and activities where conformance is critical or important. Nonconformance measures are based on detecting, enumerating, and classifying instances of nonconformance.

Although nonconformance measures are more common, conformance measures , which measure the number or fraction of activities, process variables, or outputs that comply with mandates, are an alternative. Either perspective will be effective as long as the choice is clearly defined and used consistently.

Not all nonconformances are equal. Severity levels may be defined to help delivery, support, and management personnel assign the appropriate level of priority in response and recovery efforts prompted by nonconformances. Severity levels should enable a focus on minimizing losses to the business processes supported by the services affected by the nonconformance. Typically, impact and urgency are factors in determining severity. Impact

may be measured as the amount of damage or loss incurred by the service provider or the customer as a consequence of a nonconformance. Urgency may be measured as the amount of time available between the detection of a nonconformance and the onset of the resultant impact. Nonconformances with immediate impact may be classified as most urgent, whereas those with latent or delayed impact may be considered less urgent along a sliding scale (although they may be important, even if not urgent).

A useful resource for building operational definitions of nonconformance measures is "Software Quality Measurement" [Florac 1992].

4.5.4. Performance/Satisfaction

Performance/satisfaction measures focus on the efficiency and effectiveness with which work is being completed.

Results or outcomes to be achieved by a given process or Practice are usually the focus of these measures. They are often directly related to or linked to external values and commitments made to clients. Therefore measures in this category are much more diverse and specific in nature, depending on particular contracts or agreements (including service level agreements (SLAs)), customer segments, and service definitions. These measures, however, relate to processes that are not only more visible to clients, but also closely involve or affect them. Satisfaction measures, for example, often require information gathered from clients and end users of services.

Performance measures used to support the management of service delivery and support processes typically fall into two tandem categories: efficiency and effectiveness. Measures of effectiveness often reflect a customer-centric view of the service, focusing on favored values and outcomes—in other words, desired effects. A process or Practice may be deemed ineffective if it does not produce the desired effect, regardless of how efficiently it may have been executed or implemented. Efficiency measures, however, are required to ensure that the magnitude of the desired effect is increased while decreasing, or keeping constant, the costs of producing that effect. Efficiency measures are typically ratio measures formed by dividing the output of activities in an interval by the quantity of resources used to deliver those activities.

Not all performance can be directly measured or quantified in terms of data collected from the work products of processes or Practices. Further, it is not uncommon to find service providers puzzled that customers or service personnel are unsatisfied after SLA targets have been successfully met. In such instances, the quantitative performance measures have, most likely, either been misleading or have provided too narrow a perspective to be useful. Customer satisfaction measures may therefore be required in addition to the usual set of performance measures. While satisfaction may be experienced and expressed in qualitative terms, a well-designed numerical index can help translate such feedback into quantitative information useful for analysis and communication. These indices are often required by contractual obligations and are tied to financial terms and conditions. Even when they are not, such tools for measuring satisfaction are useful in validating the views presented by other types of measures.

These indices, however, need to be periodically reviewed to ensure that they accurately model satisfaction where there have been perceptible or significant changes in customer or employee expectations, attitudes, or perceptions, as often occurs after organizational changes such as mergers, acquisitions, transition of employees (e.g. as part of sourcing deals), and changes in key personnel. Not doing so would introduce distortion between the levels of satisfaction actually expressed and the levels perceived by service provider management.

Useful resources for building operational definitions of performance and satisfaction include Sections 4.1 and 4.2 on customer and end user satisfaction in the COPC Gold Standard [COPC].

5. Implement Measures

The third step in the measurement roadmap, Implement Measures, involves data collection, analysis, and decision making. This section addresses those activities, then describes the efforts specified in the measurement Practices at each Capability Level.

5.1. Collect and Process Data

Once operational definitions are established, data need to be collected according to those definitions. For most measurement systems, that means establishing a regular and repeatable system that yields the same set of measurement points at every collection interval. Further, it requires that there be a method for storing the data over time so that they are not lost or altered. As one of the primary goals of a measurement system is the development of a long-term knowledge base, data storage should be addressed early in the system design. While spreadsheets can be used for small systems, they are inadequate from several standpoints: they do not support traditional queries well, they are susceptible to inadvertent modification, and they are typically limited in the number of entries they can contain. Because of this, a database of some sort is usually preferable as a data repository. Update access should be limited to those people who must make changes; all others should have read-only access.

The physical act of gathering data is both time-consuming and error-prone. Every reasonable effort should be made to reduce the opportunities for problems to arise. Automation can reduce both effects, and it usually has the added advantages of being timely, consistent, and repeatable. Automation is not foolproof, and checks need to be implemented in any automated measurement system, but an automated solution will greatly decrease the number of errors in the data. Access to the raw data directly from the database is the most desirable solution, as this eliminates the need for manual interaction with the data. If such access is not possible and manual input is required, input screens that perform validation should be implemented. Finally, even with automated systems, periodic reviews are necessary to ensure that the environment has not changed since the definitions were implemented. This is a common error source, and one that can lead to long periods of undetected errors that result in the potential for bad decisions.

Data should be both verified and validated before they are used. While these terms may sound similar, they refer to very different aspects of data checking. Verification is the process of insuring that the data are being collected correctly. This involves making sure the data are collected according to the operational definition; in other words, have they been collected in the manner in which we expected? Were the values collected at the correct time, from the correct place, on the appropriate intervals, etc? For automated data collection, intensive verification efforts should take place when the automation is designed, developed, tested, and implemented. Ongoing verification usually consists of ensuring that the automation continues to run as it was designed.

Validation involves making sure that the data that have been collected are correct. In other words, are the values sensible, are the data complete, have there been transcription

errors, etc. In many measurement systems this ends up being a manual effort, often only performed too late when someone views a report and asks, "is this number really right?" For large systems, manual validation is not feasible. Unfortunately, that means that validation seldom gets done in practice. While no validation method is foolproof, a great deal of it can be automated using logic that tests for outliers, null values, and the like. Such tests will always be susceptible to false positives and false negatives, but a well-designed automation system will catch many more data problems than a manual system. Sampling can also be used to periodically check data values with the original sources as a form of audit.

5.2. Types of Measurement Analysis

It is now common to find measures represented in near real-time on dashboards or other tools that serve a similar purpose for management control. While such tools can be useful, there is a very real danger that normal variability in processes will be interpreted as significant change, thereby driving unnecessary and destabilizing actions. Any real-time monitoring system should have some method for distinguishing statistically significant change from normal random variation present in processes that are in control. Statistical process control techniques such as control charts may be useful for the purpose of discerning the underlying cause of variation.

If the number of nonconformances of a specific type crosses a given threshold or control limit, it may indicate that the associated process or procedure is out of control. Observing the trend in number of nonconformances across two or more periods may validate this initial assumption. It may then be determined whether the nonconformances fit a pattern that betrays an adverse condition such as the ineffectiveness of processes or procedures, or poor conformance on part of service personnel or customers. Further analysis may reveal the condition or change that is causing the nonconformances to occur. Monitoring nonconformances is also required to review the effectiveness of actions or changes implemented earlier to eliminate their cause.

The following types of analysis are useful to interpret measurement data in service planning and delivery:

Trend analysis of performance measures, or plan versus actual measures over a period of time, will help the organization to judge whether the particular activity is improving, stable, or deteriorating. For example, a downward trend of resource utilization variance would mean improvement, an upward trend would mean deterioration. Such a simple analysis can be used for a single engagement. For an organization-level analysis, multiple trend lines of similar types of engagements can be shown on a single chart.

Control charts are useful for monitoring service-delivery processes. For example, conformance measures can be plotted on control charts. Upper and lower control limits can also be plotted on control charts to show normal process variation, and they can be compared with acceptable process variance, which is typically derived from an organization's process-capability baseline or agreed upon in the contract. Any data point

outside the control limits is a candidate for causal analysis. Processes with normal variation that exceeds acceptable variation are candidates for reengineering.

Pareto analysis is a technique that embodies the 80/20 rule (i.e., 80% of the problems are in 20% of the components). A Pareto chart is created by sorting a histogram of measurements from the largest to smallest values. From the Pareto chart, improvement techniques such as root cause analysis can be used to identify contributing factors to a performance level.

The process for organizational baseline analysis is designed to explore data for consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data. The ultimate goal is prediction. The process consists of three stages: (1) initial exploration, (2) model building or pattern identification (including validation), and (3) application (using the model with new data in order to generate predictions).

Stage 1: Exploration

This stage usually starts with data preparation which may involve data cleaning, data transformations, selecting subsets of records and—in case of data sets with large numbers of variables—performing some preliminary operations to bring the number of variables to a manageable range. Then this stage involves the identification of the most relevant variables, and the determination of the complexity and/or the general nature of models that can be taken into account in the next stage. Depending on the nature of the analysis, this may involve a simple choice of straightforward predictors for a regression model, an elaborate exploratory analyses using a wide variety of graphical and statistical methods, or an intermediate process.

The reason for the heavy reliance on graphics is that, by its very nature, the main role of organizational performance analysis is to open-mindedly explore the service provider's capability and graphics give the analysts power to do so. In combination with the natural pattern-recognition capabilities that we all possess, graphics provide unparalleled power to gain new insight into the data. The particular graphical techniques employed in organizational baseline analysis are often quite simple, consisting of various techniques:

- plotting the raw data, such as scatter plots, histograms, and time series
- plotting simple statistics, such as mean plots (used to see if the mean varies between different groups of data) and box plots (used to detect and illustrate location and variation changes between different groups of data).
- positioning such plots to maximize our natural pattern-recognition abilities, such as using multiple plots per page

Stage 2: Model building and validation

This stage involves considering various models and choosing the best one based on its predictive performance (i.e., explaining the variability in question and producing stable results across samples). There are a variety of techniques developed to achieve that goal, many of which are based on so-called "competitive evaluation of models," that is, applying

different models to the same data set and then comparing their performance to choose the best.

Stage 3: Application
The final stage involves applying the selected model to new data in order to generate predictions or estimates of the expected outcome. The models can also be used to evaluate improvement opportunities.

5.3. Measurement at Capability Level 1

Measurement initiatives in Level 1 organizations are often ad hoc, driven by the immediate business needs of the service provider and its clients. They have significant gaps in data collection and analysis in some areas, although some higher-level Practices may have been implemented.

5.4. Measurement at Capability Level 2

At Level 2, a service provider is expected to define measurable engagement objectives, define the measures needed to track its progress, and verify its success. Level 2 measurement therefore addresses the fundamental management issues of status, progress, cost, effort, and conformance. Where measurement at Level 1 may have been ad hoc, measurement at Level 2 provides a reasonably comprehensive picture of engagement performance.

Capability Level 2 is focused on consistently meeting requirements. Measurements at this Level are focused on monitoring performance against client commitments and other requirements. There are four Practices at Capability Level 2 that specifically outline how to begin addressing measurement: prf01, "Engagement objectives," del04, "Verify service commitments," prf02, "Verify processes," and knw08, "Resource consumption."

These Practices take the organization through the tasks of establishing measurable engagement-level objectives, verifying that service commitments are being met, verifying that process commitments are being met (note that these commitments may be internal), and collecting basic data on resources consumed to feed cost and estimating analysis.

Collection and analysis of this information on an engagement level helps the organization understand current performance with respect to the engagement commitments, and serves as the basis for managing improvements to performance on an engagement. Management at the engagement level using commitment, process compliance, and resource utilization measurements will provide the service provider with repeatable, predictable results for the services provided in that environment. This is a significant accomplishment that differentiates an organization at Capability Level 2 from its competitors.

Once the service provider is effectively managing its performance against commitments at the engagement level, it can begin to analyze trends over time and perform comparisons across engagements to facilitate organization-wide improvements. This sets the service provider on the path to Capability Level 3 performance.

Using the four measurement-related Practices at Capability Level 2 as a foundation, as illustrated in Figure 3, the following sections outline the key components of a measurement program for an IT-enabled service provider. Each section begins with a Practice and its rationale, followed by a measurement-oriented discussion.

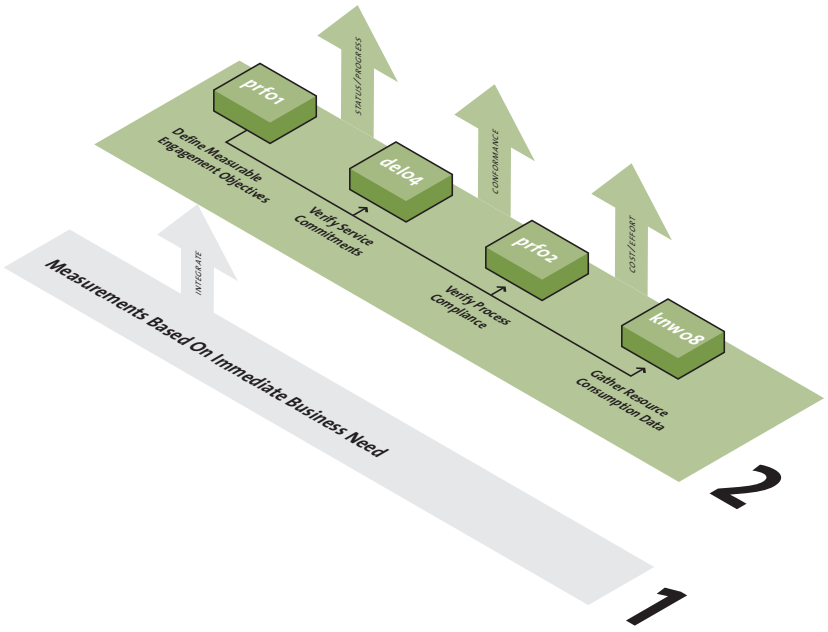

Figure 3
Measurement Practices at Level 2

5.4.1. Engagement Objectives – prfo1

> ***Define, communicate, and track engagement objectives.***
> Gain a clear understanding of the objectives for each engagement in order to enable the effective formation, management, and expansion of successful sourcing relationships. Clearly defined engagement objectives help the organization set expectations with internal and external stakeholders.
>
> Objectives should be as measurable as possible so that progress can be tracked. Objectives are typically based either on client relationship factors (e.g., increase client satisfaction by x%) or on operational factors (e.g., improve productivity by y%).

Measurement drives behavior. Therefore, a successful measurement program requires the definition and communication of engagement objectives and the alignment of what is being done and measured with those objectives. Objectives must be specified at enough detail and clarity to be measurable, allowing the organization to set targets and measure progress toward achieving its objectives. For example, at the engagement level, there may be an objective to improve the quality of service delivered. It is not until a measurable objective is specified, such as "improve first time call resolution by 5%," that the objective can be used in actionable plans and can be evaluated as to whether it is being achieved. Once the objectives are defined and it is clear how they will be measured (what data will be collected and analyzed), the other components of an operational definition for the measures should be addressed: measurement owner, data collection frequency, data analysis, and reporting.

Defining tangible engagement objectives sets the foundation for a data-oriented approach to engagement management. Capability Level 2 organizations focus on using data to understand where they are and to monitor the effects of changes (e.g., increased resources, scope and schedule adjustments) to improve overall performance against commitments.

5.4.2. Verify Service Commitments – del04

> ***Establish and implement procedures to measure and verify that service commitments are being met.***
> Verify that the organization's commitments are being met and take appropriate action when commitments are in jeopardy of being missed in order to meet service commitments and increase the client's satisfaction. Use the current performance measurements to evaluate whether the commitments are being satisfied, and identify the problems for which corrective action is needed.

The service provider has made service commitments to the client through the contract, the service level agreement and/or the Service Specification. In this Practice, the service provider focuses on outlining how its performance against those commitments is measured and reported. For each service commitment it is essential that the service provider and the client agree how the achievement of that commitment will be proven (i.e., measured). The data to be collected must be identified: what data, who will provide and collect them, how are they to be collected, and when/how often will they be collected and reported. A template for establishing operational definitions is included in Appendix C, "Operational Definition Template and Sample Measurements." These measures will largely be service-level status and/or progress indicators, but may also address service performance and various customer satisfaction indicators.

5.4.3. Verify Processes - prf02

> ***Establish and implement procedures to verify that processes are consistently performed as defined.***
> Verify that work complies with the organization's policies, procedures, and standards, and that the required work products are being produced, in order to help ensure that the work is being consistently and adequately performed. Verification is typically performed through reviews and audits by management or by personnel who are independent from those performing the work. Effective verification ensures that process requirements are being met. Process requirements may come from policies, procedures, standards, client requirements, and statutes or regulations; they may also include work product requirements either directly or by reference.

Capability Level 2 primarily focuses on engagement-level consistency; the service provider consistently meets the requirements of the client. A key tool to achieving consistency of results is consistency in the process used to deliver those results. It is this presumption that makes prf02, "Verify processes," such an important Practice. Implementation of the eSCM-SP at Capability Level 2 results in the service provider establishing and implementing defined processes. To monitor the effectiveness of those processes and to enable continual improvement, the service provider needs to have information about the use and the performance/results of those processes. This allows the service provider to address the cause of a problem rather than the symptom.

Practice del04, "Verify service commitments," gives the service provider basic information about whether or not the processes are delivering the desired (committed) result. In this Practice the service provider concentrates on measuring the use of processes. In order to confidently relate the performance of a process to its results, the organization needs to understand the use of the process. As processes for planning and delivering service are defined and implemented, it is essential to outline how the consistency and proper use of the process will be measured. Data collection must cover the steps and deliverables mandated in a process, the results of process audits, and the identified instances of nonconformance.

Using the data collected in this Practice, the service provider can manage the results of a process more effectively. For example, if the quality level of application code delivered to the client does not meet requirements (as measured by those measures identified in del04, "Verify service commitments"), the service provider must first determine whether the standard build-and-test processes were being followed. Are the standard deliverables being produced? Do process audits demonstrate that process steps are being properly performed? If the answer to these questions—as provided by the measurements—is no, then the service provider must focus on process conformance rather than on altering the process. If, on the other hand, measurements indicate that the standard process is being followed, then the service provider must focus its improvement initiatives on changing the delivery processes to improve its results.

5.4.4. Resource Consumption – knw08

> ***Establish and implement procedures to analyze and use information on resources consumed.***
> Understand resource usage for current client engagements in order to better understand and control resource utilization. This understanding is the basis of accurate estimates and planning for the resources required for new client engagements. Resource analysis enables the organization to accurately estimate future usage and balance needs, while delivering services to existing clients.

Understanding the cost drivers for the delivery of service, and managing cost/effort on an engagement is fundamental to the success of a service provider. This eSCM-SP Practice helps the service provider identify the basic data required to estimate and manage its cost and effort. Resources consumed by a service provider on an engagement may include labor, computer resources, supplies, and communications resources. All of these may be major cost drivers for the engagement and are likely to be critical to the delivery of service. Collecting and analyzing data on the consumed resources allow the service provider to plan and estimate its resource needs and usage, and manage the associated costs. Further, knowledge gained from existing engagements can be used to generate estimates for new business in the future.

5.4.5. Summary of Level 2 Measurement

A service provider at Capability Level 2 is able to consistently meet client requirements (as measured against its service commitments), achieve engagement objectives (as measured against its defined objectives), progress toward consistency of results through

the consistent application of processes (as measured by conformance measures), and understand and manage costs using information on resources consumed. The key to measurement at Level 2 is understanding the engagement objectives and commitments at a specific, measurable level; defining how those measures will be obtained, analyzed, and acted upon; and using that fundamental data for process and performance improvement.

5.5. Measurement at Capability Level 3

At Level 3, measurement is done across multiple engagements and directly addresses organizational issues, including efficiency and performance. Standardization of service delivery and commoditizing the infrastructure support improved efficiency. Performance is measured relative to organizational objectives, although most of the data are collected at the engagement level and aggregated for use at the organizational level.

Organizational learning is a key theme of the organization at Level 3. The service provider integrates the best practices followed through different engagements into assets available across the organization. Common processes provide a basis for consistent performance, and common measures provide a basis for management by fact and learning.

Through every engagement the service provider looks at cost/effort, status/progress, nonconformance, and performance/satisfaction measures to understand its performance and to initiate corrective action when required. The emphasis for organizations at Level 2 is more reactive, focusing on engagement performance management. Organizations at Level 3 focus on being proactive, focusing on organization performance management, and learning from experience for future work. Performance trends are used to control current performance, identify opportunities for improvements, and predict future performance.

At Level 3, the cost/effort, status/progress, nonconformance, and performance/satisfaction measures are still used to monitor and control engagements by the functional managers responsible for those engagements. At Level 3, performance means both business performance (e.g., market share, revenue, profitability) and operational performance (e.g., responsiveness, productivity) of the organization. Most of these measures are derived from the base measures collected at Level 2. For example, effort data and volume of service delivered data collected at Level 2 from different engagements are collated to compute the units of service delivered per FTE, which is one indicator of the operational efficiency of the organization.

As shown in Figure 4, the core measurement Practices at Level 3 are prf04, "Organizational objectives," prf05, "Review organizational performance," and knw04, "Process assets." Performance targets are set as part of the organizational objectives and the set of measures to be used in analyzing organizational performance are determined in prf04. Organizational performance targets will largely be derived from analyzing engagement level performance as described for Level 2. The periodic review and analysis of the performance measures are addressed in prf05. This is done to verify that the organizational objectives are being met and to address identifying opportunities for improvement. A measurement repository for the organization is established in knw04.

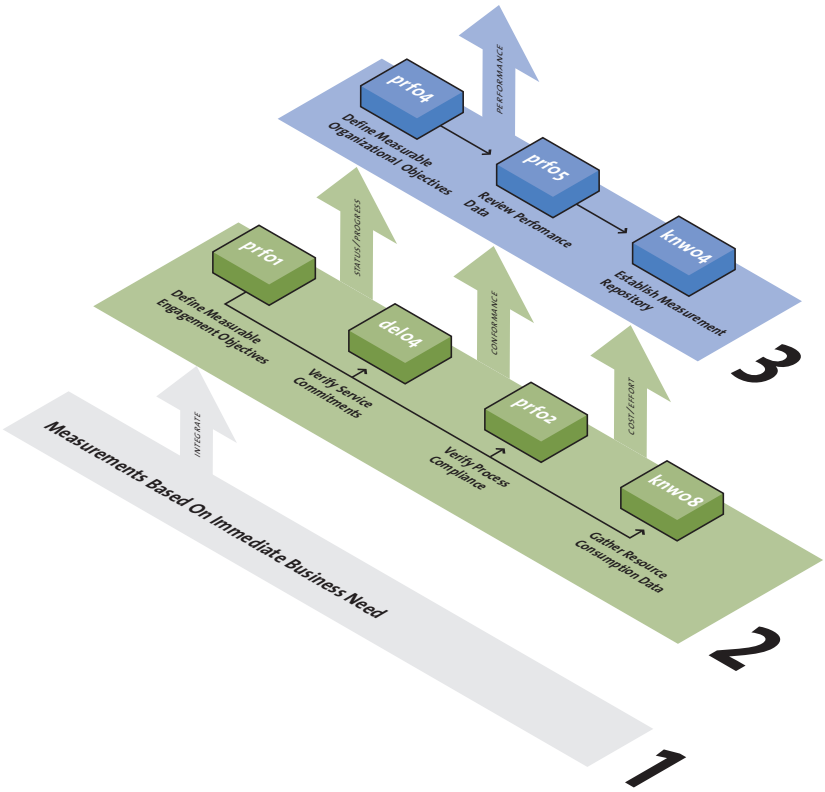

Figure 4
Measurement Practices at Level 3

5.5.1. Organizational Objectives – prf04

> ***Define, communicate, and track organizational objectives.***
> Have a clear understanding of the organization's objectives in order to establish the organization's long-term direction and enable the effective formation, management, and expansion of successful sourcing relationships. Clearly defined and measurable organizational objectives enable management to set and attain performance expectations for the organization. These objectives are usually long term and are derived by analyzing the organization's environment (including client expectations, competitors, and business trends). To be effective, organizational objectives must be measurable, address the needs of stakeholders, and establish both desired and expected performance targets.

At Level 3, the organization defines its objectives, which become the drivers of every other activity performed by the organization. These objectives typically define both short- and long-term goals of the organization, and should be aligned with the vision and/or long-term business objectives of the organization. At Capability Level 3, individual engagements will still define their respective engagement objectives, but those objectives should be in line with organizational objectives.

From the measurement perspective, the definition of organizational objectives is of special significance because it helps to establish the organization's performance management framework.

Analysis of customer expectations, current and past performance data, market trends, etc., provide input for defining organizational objectives. Most of the data required for analysis are collected at Level 2 (e.g., past performance data from engagements, client

and prospective client expectations, client feedback), though additional data may need to be collected at Level 3. Two Level 3 Practices (knw05, "Engagement knowledge" and rel05, "Stakeholder information") facilitate analysis of data. At the organization level, variation in performance measures is expected. Data collected from similar types of engagements should be consolidated and plotted in a scatter chart. The out-of-range data points should be analyzed for the causes of variation. Organization-level data analysis builds the foundation for determining the organization's capability baselines and benchmarking at Level 4.

Organizational objectives set measurable performance targets for the organization. They should include both short-term and long-term performance targets (e.g., reducing time to market by x% within the next 12 months and growing market share by y% by end of year Y). After setting organization level performance targets, key processes are identified that contribute to achieving performance targets and setting improvement goals. While setting performance targets for individual processes, organizations need to be sensitive to current process capabilities and set realistic performance targets. For example, a 40% reduction in first-call resolution time may be beyond the capacity of the existing call resolution process. If a quantum jump in performance is required, the organization needs to undertake special improvement programs (e.g., Six Sigma) and/or adopt breakthrough innovations. These types of changes are supported by Level 4 capabilities.

After defining its measurable performance goals, the organization needs to establish methods to verify the achievement of its performance targets. Without an appropriate performance management system, the whole exercise of objective and goal setting could be futile. The performance management system includes instructions on performing the following activities:

- Identify the attributes to measure for tracking achievement of the organizational objectives.
- Determine the operational definitions of the performance measures for those attributes.
- Identify the personnel responsible for data collection and analysis.
- Provide guidance on how to interpret data.
- Determine an acceptable performance range for each measure.
- Define a corrective action procedure.

The use of tools such as the Balanced Scorecard is not mandatory, but it is almost imperative.

5.5.2. Review Organizational Performance – prf05

> ***Establish and implement procedures to review organizational performance.***
> Analyze key performance measurements and base management decisions on the results in order to manage and improve organizational performance. Using well-defined performance measurements as a basis for making managerial decisions helps to ensure the consistency and relevance of the decisions. Periodic reviews allow management to compare performance to the performance requirements, thereby enabling the organization to make performance improvements.

After establishing an organization performance management system in the previous Practice, this one implements that system. This Practice also defines additional parameters needed to complete the performance management system:

- sources of data required for performance review
- periodicity of performance review

The focus of this Practice is on collecting the identified performance measures, comparing them with the expected performance, and taking corrective actions when the performance deviates from the expected result. The primary objective of this practice is to track progress toward achieving the stated organizational objective.

A performance review should also consider the effectiveness and efficiency of the processes. For example, an organization that has set a target to reduce its first-call resolution time by x% may have achieved this target, but if it is found that its customer satisfaction score has decreased, then the call resolution process is still not effective.

Performance reviews take place at different levels in an organization. The mechanism of performance review, the periodicity, and the types of information used for the reviews will vary, depending on the organization's culture and its business needs. Typically, middle management uses more measures and details associated with the analysis, whereas top management looks at fewer key performance indicators.

5.5.3. Process Assets – knw04

> ***Establish and maintain a set of process assets for use across the organization.***
> Create a consistent set of process assets across engagements in order to help the organization achieve economies of scale and apply best practices from prior engagements to similar new engagements. Focusing on systematically collecting knowledge is fundamental to organizational learning. Providing a coordinated set of process assets enables the collected knowledge to be made available consistently across engagements.

This Practice addresses the requirements of establishing and maintaining a set of process assets for the organization, which include process performance data. Performance data are useful for management (as described in the previous sections), estimation, and forecasting.

At Level 3, the process assets repository typically contains data analysis results (section 5.4.1), and performance review results (section 5.4.2), etc. At Level 4, it contains capability baselines, benchmarking data, etc. The design and content of the repository must be consistent with the information needs at various levels (section 5.4.2) of the organization.

This Practice is very closely linked with prf03, "Adequate resources," which addresses the need for establishing and implementing the knowledge system of the organization. A measurement repository is an important part of an organization's knowledge system.

5.5.4. Summary of Level 3 Measurement

Measurement at Level 3 is performed at the organization level rather than the engagement level. This permits Level 3 organizations to actively engage in organizational learning. They no longer perform in a reactive manner, but, using performance trends and lessons learned across engagements, they proactively focus on predicting future performance.

5.6. Measurement at Capability Level 4

By the time a service provider is ready to consider measurement from a Level 4 perspective, its measurement system has already demonstrated a considerable level of maturity. The organization consistently meets requirements, has demonstrated consistency across engagements, and has the historical data to verify its performance. At this point, it is time to expand the measurement focus from consistency to excellence. As illustrated in Figure 5, the measurement-focused Practices at Level 4, are prfo8, "Capability baselines," and prfo9, "Benchmarking." These center around benchmarking, which Robert Camp defines as "research or information gathering that allows a manager to compare his or her function's performance to the performances of the same functions in other [organizations]" [Camp 1989]. Clearly, before an organization can take advantage of benchmarking it must thoroughly understand its own operation, thus capability baselines (prfo8) are key to the success of any Level 4 measurement program.

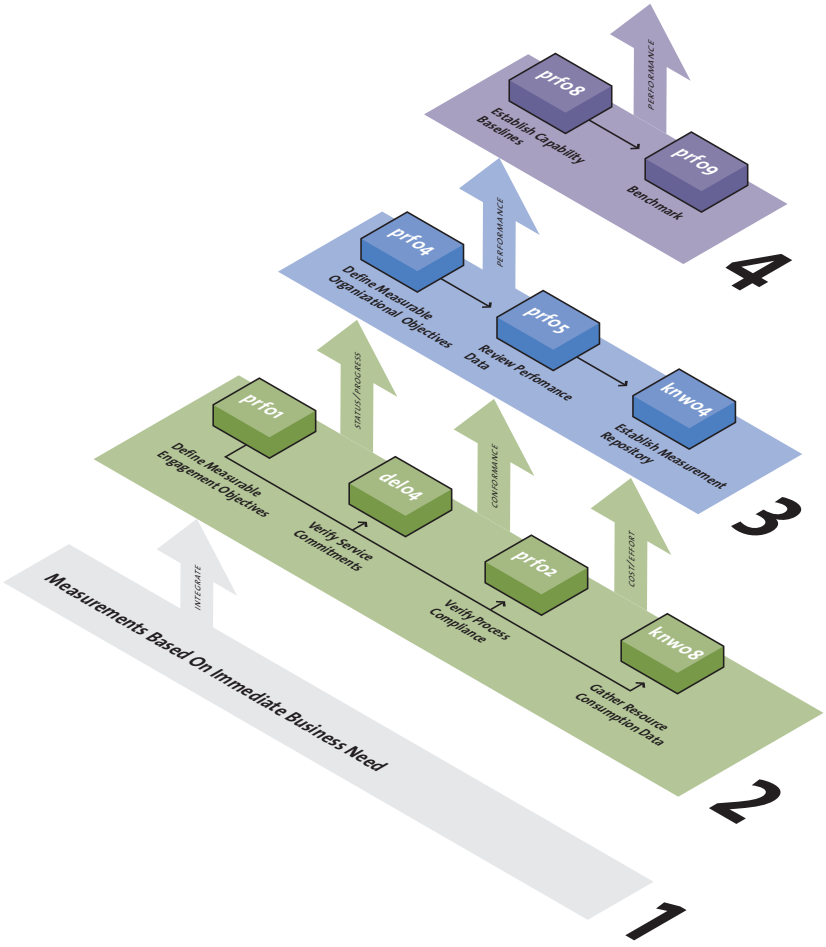

Figure 5
Measurement Practices at Level 4

5.6.1. Capability Baselines – prf08

Define capability baselines for the organization by analyzing performance data.
Use well-defined capability baselines in order to predict performance, make accurate estimates, and improve the quality of services. A statistically significant capability baseline provides a basis for the organization to analyze whether performance deviations are within expected ranges or if they represent exceptions that need to be investigated and addressed. Doing this provides a firm basis for deciding how to allocate resources and make organizational improvements.

Measurement in prf08 builds on all of the accomplishments from Level 3. Here the focus is on moving beyond the context provided by the historical data that have been collected across the organization to characterizing the nature of the processes the data represent. This characterization is most often accomplished through the use of statistical process control (SPC) techniques. One of the leading experts in the field, Donald Wheeler, calls SPC charts "process behavior charts." He does this intentionally to draw attention to the real purpose of the charts: to show graphically the behavioral characteristics of a given process. Control charts do not control processes, and they do not represent specifications or requirements in any way. They are a powerful tool to aid in understanding whether a process is predictable (within a range) and to monitor whether that predictable behavior changes over time. Once predictability is established, SPC becomes the basis for evaluating an organization's improvement over time.

Capability measurements are generally necessary at two levels: across the organization as a whole, and at the individual engagement level. Virtually all of the necessary data are available in the measurement repository created in Level 3 (knw04, "Process assets"); the focus here is on using the data as a whole. While organization-wide measurements are necessary to understand overall capability, engagement-level measurements are also needed to ensure that capability is consistent across the engagements. Using SPC on aggregated data alone is a risky proposition, and maintaining engagement-level measures protects against missed opportunities. Further, as the organization seeks to improve its overall capability it will necessarily seek improvements in specific areas, and engagement-level measurements—being more sensitive—will be the first to indicate that progress is being made.

As requirements and service levels often vary across engagements, care must be taken when comparing baselines across engagements to ensure that one is comparing similar environments. Often, differences in baselines are a result of differences in service levels, so the organization must normalize baselines before making value judgments around relative performance levels.

5.6.2. Benchmark – prf09

Benchmark organizational performance to identify opportunities for improvement.
Benchmark organizational practices using best practices from reference models, competitors, or industry leaders in order to identify opportunities for improvement. Benchmarking allows the organization to objectively analyze the processes it uses to deliver its service. When compared with the organization's goals and objectives, these analyses provide a basis for deciding how to allocate resources in order to make organizational improvements.

A high-quality measurement system is a precursor to any benchmarking activities. Even with good data, benchmarking is a complex task requiring careful attention to detail, as the act of comparing two or more environments requires an organization to account for the many structural differences that drive real differences in outcomes. While measurements are important to benchmarking, benchmarking is by no means a strictly quantitative endeavor. To be effective, benchmarking requires an organization to learn not only the magnitude of the gap that exists between where they are and where they want to go, but also the best-of-breed practices that will get them there. Thus, process analysis and comparison is a significant part of a successful benchmarking activity.

Measures such as productivity, responsiveness, reliability, and throughput are just some of the key items that are considered in benchmarks, and all of these need to be well understood within the organization before embarking on a benchmark. As seen above, capability baselines give an organization an accurate view of their own performance, and therefore provide the basis for a benchmarking study. Most of the measures an organization needs for a benchmark will already be part of the measurement repository, but some new measures are likely to be needed in most benchmarking studies. If these new measures prove valuable to the organization, they should become part of the ongoing measurement system.

Establishing the methods to be used in a benchmarking project is at least as important as the actual measurements themselves. Once the "area" to be benchmarked is chosen, it is critical to identify the measurements that will be used to perform the comparisons as well as the methods that will be used to judge the comparisons. It is important to determine the criteria for judging the results before the data are collected; deciding after the data have been collected opens one to charges of biased analysis. As with capability baselines above, it is also critical here to account for the natural variability in the benchmark measurements when one performs comparisons. Statistical techniques such as SPC, Analysis of Means (ANOM), Analysis of Variance (ANOVA), and hypothesis testing are all appropriate techniques under the right conditions. Simply comparing two numeric results is never appropriate, regardless of how often one sees this approach in practice.

Finally, because no two environments are ever identical, it is almost always necessary to adjust the data for the underlying environmental differences—a process often referred to as normalizing. Issues such as size, complexity, service levels, and government regulations can often have a significant effect on the measurements under study. If these differences are not accounted for, one can easily draw the wrong conclusions. For example, if one were attempting to understand the efficiency of two groups performing a certain manual operation, and group one completed it in five hours and group two completed it in eight hours, one might be tempted to conclude that group one was more efficient. If, however, we are told that group one used two people in parallel and group two used only one person, the conclusion reverses. In this example, elapsed time is not a sufficient measure of efficiency for comparing these two groups; some adjustment must be made to account for the resource usage.

5.6.3. Summary of Level 4 Measurement

Service providers at Capability Level 4 are focused on continual improvement. This focus requires both an understanding of the organization's current capabilities and an awareness of the organization's capability vis-à-vis others performing similar tasks. Capability baselines established through statistical means such as SPC position an organization to accomplish both tasks.

5.7. Measurement at Capability Level 5

Service providers at Capability Level 5 demonstrate measurable, sustained, and consistent performance excellence throughout the organization over time. While there may be some question about the effect of an innovation in the short term for a Level 4 organization, the Level 5 organization can demonstrate the long-term effect of multiple innovations as measurable improvements in performance.

5.8. Measurement Activities by Phase and Level

From an eSCM-SP perspective, the measurement-related Activities performed during the phases and for each Capability Level are summarized in Table 2. For each phase and Capability Level, performance monitoring parameters must be built into the sourcing agreement and assessed on an on-going basis. These measures form an end-to-end measurement system embodied in a measurement architecture that addresses all critical issues linking technology organization performance, key contract parameters, and business performance.

Table 2
Key Measurement Activities by Sourcing, Phase, and Capability Level

Level/Phase	Level 1	Level 2	Level 3	Level 4	Level 5
Initiation		Baseline current performance. Estimate costs, staffing, and technology performance.	Build estimation model, tracking model, and technology performance and capacity model.	Implement knowledge sharing and optimization.	Optimize models.
Delivery	Contract specific measures.	Track and control engagement and supplier/ partner performance.	Measure and improve organizational performance.	Implement performance engineering.	Maintain continual improvement.
Completion		Perform root-cause analysis. Account for resources transferred back.	Institute organizational measurement. Maintain service continuity.	Transfer knowledge.	
Ongoing	Profit and performance.	Maintain engagement-by-engagement performance and status. Perform engagement resources and conformance measurements.	Maintain consistency. Reduce variation from engagement to engagement. Achieve internal business objectives.	Benchmark internal capability baselines and external businesses. Seek world-class innovation.	Maintain continual improvement.

6. Measure the Impact of the eSCM-SP

Decisions about adoption of the eSCM-SP will be driven by a variety of factors including client requirements, adoption of the eSCM-SP by direct competitors, and the value the organization derives from Model adoption. This section provides a brief discussion of how to measure value. Analyzing how the eSCM-SP increases value will enable adopters to make informed business decisions about their improvement paths and provide convincing arguments to sponsors and clients about the merits of implementing the eSCM-SP. This section is not intended to be a comprehensive discussion on measuring value, or to mandate a specific way to measure the value of the eSCM-SP. It is intended to acquaint service providers with some of the subtleties and complexities of this type of analysis.

Defining value measures, determining the appropriate methods and tools to analyze these measures, and determining the appropriate level of analysis (i.e., firm-level, project-level, client-level) are a few of the more difficult issues to consider before starting an analysis. There are many books and articles available on measuring value, including those from the disciplines of finance [Brealey 2003, Copeland 1996, Stern 2001], accounting [Maher 1997], and software engineering [Reifer 2001, Emam 2003, Erdogmus 2004, Solingen 2004].

Value itself can be represented in various ways, but the ultimate measure of value is increased economic profitability. Economic profitability is not the same as accounting profitability. Accounting profits are based on complex rules and regulations which will differ from country to country. The rules used for determining accounting profits also tend to understate economic value since they are designed to expose downside risk to lenders. Calculating accounting profit provides a conservative estimate of liquidation value, but does not provide a good estimation of market value, what an investment is worth to a firm owner, such as a shareholder. By looking at economic profitability you can compare investments across projects (which may cross geographic boundaries), and time (accounting for differences in project length, and risk).

6.1. Economic Profitability

The best way to determine economic profitability is to look at the cash flows for a project, or set of projects. The cash flows need to include positive (e.g., increased revenues) and negative (e.g., personnel and hardware costs) cash flows of the project. To account for the fact that money received today is not worth the same as money received tomorrow, the cash flows need to be discounted over the time period of the project. This type of analysis is called Net Present Value (NPV), or Discounted Cash Flows:

$$NPV = \sum_{t=0}^{n} \frac{CF_t}{(1+d)}$$

where

- CF_t = Net cash flows (positive or negative) at the end of time period t
- d = Discount rate for future net cash flows
- t = Time period when the net cash flow occurs
- n = Life of the project, in years

The calculation of this formula is fairly straightforward. The difficulty comes in determining the cash flows, discount rate, and time period. For cash flows, the issue of what costs and what benefits can legitimately be associated with a Practice must be carefully considered. For example, are the training costs for a new Practice included? What if the training for that Practice is embedded in a larger set of training materials (e.g., training for risk management is added to existing management training)? Are the verification costs associated with performing quality assurance for the new Practice included? In terms of benefits, how does an organization value the satisfaction (or delight) of a customer? How does the organization value the impact of a Practice on employee morale?

Many of the benefits associated with initiating a project are often intangible. To the extent possible the organization needs to define ways to translate intangible benefits (e.g., improved employee morale, reduced effort) into monetary terms. Some benefits will be easier to translate than others, e.g., hours saved from reduced effort can be converted into dollars. Considering some of the lessons learned in cost-of-quality analysis may be helpful in considering how to quantify some of the benefits [Krasner 1999, Campanella 1999].

The discount rate reflects risk associated with that project and is made up of the market demands of a return on capital investments and an inflation rate. Some projects will be riskier than others, which should be reflected in the discount rate. The discount rate can be broken into its component parts

$$(1+d)^{-t}=[(1+r)(1+i)]^{-t}$$

where

- d= Discount rate for future net cash flows
- r= Real return on capital required for future net cash flows
- i= Expected inflation rate on future net cash flows
- t= Time period when the net cash flow occurs

It is important to look over the entire life of a project to determine its value. Positive benefits, especially on complex undertakings, often accrue over a long period of time. If the time period covered by the analysis is too short, all benefits received from the project may not be accounted for in the analysis, and decisions made based on the results will be ill informed.

Due to the uncertainty of estimating the different variables in an NPV analysis, organizations will often conduct sensitivity analyses. Using sensitivity analysis an organization can see how sensitive a project's outcomes are to changes in estimates of net cash flows, timing of cash flows, and the discount rate. Another common type of analysis is the use of NPV scenarios to estimate the best, worst, and most likely project outcomes.

6.2. Other Value Criteria

There are times when an organization is interested in other value criteria besides economic profitability. These include some of the more intangible criteria like customer satisfaction and employee morale, or process-oriented criteria such as percent of service levels achieved. Collection and measurement of these criteria are often important strategic objectives, and selection of which value criteria to address should be based on the organizations strategic

planning. While these criteria are often difficult to quantify in monetary terms, they can, and should be, defined in as consistent, and measurable a way as possible to allow for analysis.

6.3. Other Types of Analyses

Organizations will often use types of analyses other than NPV to analyze value. Most of these are not recommended as your main analysis method, but can be used when the cost of conducting NPV analyses are too high, or when a quick analysis method is required. Below is a brief discussion of several commonly used analysis methods.

6.3.1. Payback

Payback refers to the amount of time it takes to return the initial investment of a project. The value assigned for payback is easy to calculate and understand. For instance, if a project's investment is $500,000, and the cash flows are $75,000 for year one, $125,000 for year two, and $200,000 for years three through five, the project will break even in the middle of year four. While easy to interpret, payback analysis should not be used as a sole tool for deciding which projects to undertake. The two main problems with this type of analysis is that it ignores the time value of money or the value of cash flows after the payback period.

6.3.2. Internal Rate of Return (IRR) and Return on Investment (ROI)

The Internal Rate of Return (IRR) is defined as the rate of return that equates the present value of the future cash flows with the cost of the investment [Pyzdek 2003]. Effectively, the IRR is the value of *d* in the NPV equation above that causes the NPV to be equal to zero. The appeal of the IRR is that it provides a number that is easy to interpret and compare. There are, however, several disadvantages to using this type of analysis.

The IRR assumes that all net cash flows can be reinvested at the project's internal rate of return. This is different from the NPV assumption that net cash flows are invested at the company's cost of capital rate. This difference in assumption can cause the outcomes of the IRR and NPV analyses to be different. For instance, if you are choosing between two projects, the IRR of Project A may be higher than Project B, but the NPV of Project B is higher than Project A. Without conducting an NPV the organization would not know that they are giving up economic profits by undertaking the project that has the higher internal rate of return.

Another problem with IRR is that when net cash flows change signs (e.g., going from positive to negative) between periods there is more than one internal rate of return. This means that every time there is a net cash flow sign change during the life of the project there will be multiple internal rates of return, one for each sign change. For a more complete discussion of the issues related to calculating internal rates of returns see [Brealey 2003] and [Maher 1997].

6.3.3. Economic Value Added

Economic Value Added (EVA) was developed by Stern Stewart & Co [Stewart 1999]. At the corporate level EVA is equal to operating profits minus the cost of all capital employed to produce those operating profits or earnings. EVA has some specific definitions that are different from standard accounting requirements to account for various categories of expenses and earnings at the corporate level and how they should be applied. The net purpose is to remove all accounting distortions from the financial analysis.

When properly executed, EVA is a forward-looking measure of value and performance for the whole corporation. Stewart points out that "...projecting and discounting the EVA to be generated by an entire company automatically sums to the net present value of all the firm's past and projected capital investment projects" (pg. 3). The use of the cost of capital in the calculation helps align expenditures more closely with future value contributions to the business by removing expenditures that could have a positive return using just ROI analysis. The parent company advances capital to its business lines at an interest rate equal to the cost of the capital, *c**.

EVA, like ROI and other accounting-based techniques, is subject to distortion caused by inflation and by the asset structure of the firm (mix of depreciable and non-depreciable assets and currency of those assets).

EVA = NOPAT – (c* x Capital investment)

ROI = EVA / Capital Investment

where

EVA= Economic Value Added
NOPAT= Net Operating Profit After Tax
c*= cost of all capital
Capital Investment= total cost of the project

6.3.4. Balanced Scorecard

Balanced Scorecard is a management and measurement system that acts as a link between a company's strategic intent and its operational measurements [Kaplan 1996]. The "balanced" aspect is that this system does not only consider a financial perspective but adds elements that contribute to a company's success. The Balanced Scorecard codifies and quantifies strategies from four perspectives: financial, internal business processes, customer, and learning and growth. The "balance" of the scorecard is achieved by attributing a weighting or percentage to each objective and summing to 100%. The weighting then helps all line management craft actions and decisions that are most consistent with the corporate strategy as driven by the weighting for each perspective and the underlying measurements. Each perspective includes objectives, measures, targets and initiatives.

The benefit of Balanced Scorecard is that it combines the management and measurement systems, helping to keep line management decision-making consistent with corporate strategy. It can be easily customized for specific situations and a number of consulting groups have published enhancements. To be effective the Balanced Scorecard must be tied to the company's compensation system. Additionally, the measurements must be consistent and tied to the groups that actually affect the measurements. The

downside is that the system may be easier to "game" because some measurements are less quantifiable. As a management system versus a technique, it requires executive commitment to implement.

6.4. Summary

Value analysis must be performed in an overall business context. In many cases, conducting an analysis of the eSCM-SP Practices and their implementations will not add value because the analysis will not affect the decisions being made. As Humphrey has pointed out, requests for a cost/benefit analysis are frequently a form of resistance since most senior management decisions are based on intuition with a seasoning of financial judgment [Humphrey 1989]. When analyses are needed, either to help decide which projects to undertake or to determine the benefits of a project that has already been implemented, the organization should analyze the discounted net cash flows to determine value. The three case studies in Appendix F provide a glimpse of how a service provider might build its capability for determining value. While they are not comprehensive examinations of each of the relevant issues, these case studies should help a service provider understand the kinds of issues to address in determining value.

7. Conclusions

The importance of measurement to effectively and efficiently implement the eSCM-SP and its Practices is undeniable. Establishing an effective measurement program is not a simple matter.

Measurement should serve the business needs of the organization, so goal-driven measurement is an appropriate philosophy. Measurement should support a balanced and comprehensive view of business needs, therefore strategies such as the those based on a Balanced Scorecard are appropriate.

Ultimately, service providers should strive for a management-by-fact style. Discussions and negotiations with stakeholders that are founded in fact—especially quantified facts—tend to be much more successful and supportive of win-win results. Continual, measurable improvement is only possible when measures support the business objectives vital to the organization.

The emphasis of the measures is expected to expand as organizations move up the eSCM-SP Levels, from special studies, to engagement, to organization, to industry. The sophistication of the analyses is also expected to increase from exploratory data analysis to statistical process control and simulation models.

Focusing on achieving the business objectives, the eSCM-SP provides a progressive measurement path through the Capability Levels, with the Practices at each Level building on the foundation established at the previous Levels. Level 1 organizations are likely to have significant gaps in capability at each Level, although some higher-level Practices may have been implemented. Measurement initiatives in Level 1 organizations are often ad hoc, driven by the immediate business needs of the service provider and its clients.

At Level 2, a service provider is expected to define measurable engagement objectives and define the measures needed to track its progress toward achieving those objectives until it can be verified that the objectives have been satisfied. Level 2 measurement therefore addresses fundamental management issues of cost/effort, status/progress, nonconformance, and performance/satisfaction. Where measurement at Level 1 may have been ad hoc, measurement at Level 2 provides a reasonably comprehensive picture of engagement performance. The Level 2 organization is expected to define measurable engagement objectives, and identify measures to verify achievement of engagement objectives. It is also expected that all activities performed be aligned with engagement objectives.

At Level 3, measurement is done across multiple engagements and directly addresses organizational issues, such as efficiency and performance. Performance is measured relative to organizational objectives, although most of the data are collected at the engagement level and aggregated for use at the organizational level. Organizational performance targets are derived from analyzing engagement-level measures established at Level 2.

At Level 4, the analysis of data for the engagement and organization is statistically sophisticated, since capability baselines and benchmarks of organizational performance

are defined. The measurement focus of Level 4 is understanding the impact of change on performance and variation since the long-term goal is to improve performance and decrease variation. In the short term, the learning curve associated with adopting innovative technologies and processes may lead to decreased performance and increased variability, so it is crucial to monitor the ultimate impact of an innovation as it is adopted and deployed.

Service providers at Capability Level 5 demonstrate measurable, sustained, and consistent performance excellence throughout the organization over time. While there may be some question about the effect of an innovation in the short term for a Level 4 organization, the Level 5 organization can demonstrate the long-term effect of multiple innovations as measurable improvements in performance.

8. References

[Austin 1996] Austin, Robert D. 1996. *Measuring and Managing Performance in Organizations.* New York, NY: Dorset House Publishing.

[Baldrige] Baldrige National Quality Award. www.quality.nist.gov.

[Basili 1996] Basili, Victor R. 1996. "Applying the Goal/Question/Metric paradigm in the experience factory." *Software Quality Assurance and Measurement: A Worldwide Perspective.* Edited by Norman Fenton, Robin Whitty, and Yoshinori Lizuka. London: International Thomson Publishing.

[Brealey 2003] Brealey, R., S. Myers, and R. Myers. 2003. *Principles of Corporate Finance.* McGraw Hill/Irwin.

[BSI 2002] BDD/3 Technical Committee. 2002. *BS 15000-1:2002, IT Service Management, Part 1: Specification for service management.* British Standards Institution.

[Camp 1989] Camp, Robert C. 1989. *Benchmarking: The Search for Industry Best Practices That Lead to Superior Performance.* Milwaukee, WI: Quality Press.

[Campanella 1999] Campanella, J. (ed). 1999. *Principles of Quality Costs: Principles, Implementation, and Use, Third Edition.* ASQ Press.

[Chang 1997] Chang, R. Y., and P. De Young. 1997. *Measuring Organizational Improvement Impact.* Irvine, CA: Richard Chang Associates.

[Chrissis 2003] Chrissis, Mary Beth, Mike Konrad, and Sandy Shrum. 2003. *CMMI: Guidelines for Process Integration and Product Improvement.* Boston, MA: Addison-Wesley.

[COPC] Customer Operations Performance Center (COPC-2000). www.copc.com.

[Copeland 1996] Copeland, T., T. Koller, J. Murrin. 1996. *Valuation: Measuring and Managing the Value of Companies.* New York, NY: John Wiley & Sons, Inc.

[Deming 1986] Deming, W. Edwards. 1986. *Out of the Crisis.* Cambridge, MA: MIT Center for Advanced Engineering Study.

[Emam 2003] Emam, K.E. 2003. *The ROI from Software Quality: An Executive Briefing.* K Sharp Technology, Inc.

[Erdogmus 2004] Erdogmus, H., J. Favaro, and W. Strigel. 2004. "Return on Investment." *IEEE Software,* vol. 21, no. 3: 18-22.

[Florac 1992] Florac, W.A. 1992. *Software Quality Measurement: A Framework for Counting Problems and Defects.* CMU/SEI-92-TR-22. Pittsburgh, PA: Carnegie Mellon University.

[Goethert 1992] Goethert, W.B., E.K. Bailey, and M.B. Busby. 1992. *Software Effort and Schedule Measurement: A Framework for Counting Staff-Hours and Reporting Schedule Information.* CMU/SEI-92-TR-21. Pittsburgh, PA: Carnegie Mellon University.

[Gopal 2002] Gopal, Anandasivam, M.S. Krishnan, Tridas Mukhopadhyay, and Dennis R. Goldenson. 2002. "Measurement Programs in Software Development: Determinants of Success." *IEEE Transactions on Software Engineering,* vol. 28, no. 9: 863-875.

[Guha 2005] Guha, Subrata, William E. Hefley, Elaine Hyder, Majid Iqbal, and Mark C. Paulk. 2005. *Comparing the eSCM-SP v2 and ISO 9001: A comparison between the eSourcing Capability Model for Service Providers v2 and ISO 9001:2000 (Quality Management Systems—Requirements).* CMU-ITSQC-05-001. Pittsburgh, PA: Carnegie Mellon University.

[Guha forthcoming] Guha, Subrata, William E. Hefley, Elaine Hyder, Majid Iqbal, and Mark C. Paulk. Forthcoming. *Comparing the eSCM-SP v2 and COPC-2000®: A comparison between the eSourcing Capability Model for Service Providers v2 and Customer Operations Performance Center (COPC)-2000 CSP Gold Standard, Release 3.4.* Pittsburgh, PA: Carnegie Mellon University.

[Haimes 1998] Haimes, Yacov Y. 1998. *Risk Modeling, Assessment, and Management*. New York, NY: Wiley.

[Humphrey 1989] Humphrey, W.S. 1989. *Managing the Software Process.* Addison-Wesley.

[Hyder 2004] Hyder, Elaine B., Keith M. Heston, and Mark C. Paulk. 2004. *The eSourcing Capability Model for Service Providers (eSCM-SP) v2, Part 1: Model Overview.* CMU-ISRI-04-113. Pittsburgh, PA: Carnegie Mellon University.

[Hyder 2004a] Hyder, Elaine B., Keith M. Heston, and Mark C. Paulk. 2004. *The eSourcing Capability Model for Service Providers (eSCM-SP) v2, Part 2: Practice Details.* CMU-ISRI-04-114. Pittsburgh, PA: Carnegie Mellon University.

[IEEE-STD-610 1990] ANSI/IEEE Std 610.12. 1990. "IEEE Standard Glossary of Software Engineering Terminology." IEEE Computer Society.

[Iqbal 2004] Iqbal, Majid, Subrata Guha, William Hefley, Elaine Hyder, and Mark Paulk. 2004. *Comparing the eSCM-SP v2 and BS 15000: A comparison between the eSourcing Capability Model for Service Providers v2 and BS 15000-1:2002 (IT Service Management).* CMU-CS-04-129b. Pittsburgh, PA: Carnegie Mellon University.

[Iqbal forthcoming] Iqbal, Majid, Subrata Guha, William E. Hefley, Elaine Hyder, and Mark C. Paulk. Forthcoming. *Comparing the eSCM-SP v2 and COBIT(R): A comparison between the eSourcing Capability Model for Service Providers v2 and Control Objectives for Information and related Technology (COBIT) 3rd Edition.* Pittsburgh, PA: Carnegie Mellon University.

[ISO 1996] ISO/IEC 14598-1:1999. 1999. "Information Technology—Software Product Evaluation—Part 1: General Overview." International Organization for Standardization and International Electrotechnical Commission..

[ISO 1998] ISO/IEC 15504-9:1998. 1998. "Information Technology—Software Process Assessment—Part 9: Vocabulary." International Organization for Standardization and International Electrotechnical Commission.

[ISO 2000] ISO 9001:2000. 2000. "Quality Management Systems—Requirements." International Organization for Standardization and International Electrotechnical Commission.

[ISO 2000a] ISO 9004:2000. 2000. "Quality Management Systems—Guidelines for Performance Improvements." International Organization for Standardization and International Electrotechnical Commission.

[ISO 2002] ISO 15939:2002. 2002. "Software Engineering—Software Measurement Process." International Organization for Standardization and International Electrotechnical Commission.

[ITGI 2000] ITGI. 2000. *Control Objectives for Information and Related Technology.* COBIT, 3rd Edition. IT Governance Institute. www.isaca.org/cobit.htm.

[Kaplan 1996] Kaplan, Robert S., and David P. Norton. 1996. *The Balanced Scorecard: Translating Strategy into Action.* Boston, MA: Harvard Business School Publishing.

[Kaydos 1999] Kaydos, Will. 1999. *Operational Performance Measurement: Increasing Total Productivity.* Boca Raton, FL: St. Lucie Press.

[Krasner 1999] Krasner, H. 1999. "The Payoff for Software Process Improvement: What It Is and How to Get It," Chapter 9 in *Elements of Software Process Assessment and Improvement*. Edited by K.E. Emam and N.H. Madhavji. IEEE Computer Society Press.

[Kumar 2001] Kumar, B., V. Mahendra, E. Hyder, E. Nawrocki, K. Madhu, and R. Gupta. 2001. *eSCM Annotated Bibliography.* CMU-CS-01-125/CMU-ISRI-01-100. Pittsburgh, PA: Carnegie Mellon University.

[Maher 1997] Maher, M. 1997. *Cost Accounting: Creating Value for Management*. Irwin.

[March 1996] March, Artemis. 1996. "A Note on Quality: The Views of Deming, Juran, and Crosby." *IEEE Engineering Management Review,* vol. 24, no. 1: 6-14.

[McGarry 2002] McGarry, John, David Card, et al. 2002. *Practical Software Measurement: Objective Information for Decision Makers*. Boston, MA: Addison-Wesley.

[Park 1996] Park, Robert E., Wolfhart B. Goethert, and William A. Florac. 1996. *Goal-Driven Software Measurement—A Guidebook*. CMU/SEI-96-HB-002. Pittsburgh, PA: Carnegie Mellon University.

[Paulk 1995] Paulk, Mark C., Charles V. Weber, Bill Curtis, and Mary Beth Chrissis. 1995. *The Capability Maturity Model: Guidelines for Improving the Software Process*. Reading, MA: Addison-Wesley.

[Paulk forthcoming] Paulk, Mark C., Subrata Guha, William E. Hefley, Elaine Hyder, and Majid Iqbal. Forthcoming. *Comparing the eSCM-SP v2 and Related Models and Standards: A Comparison Between the eSourcing Capability Model for Service Providers v2 and Related Models and Standards*. Pittsburgh, PA: Carnegie Mellon University.

[Paulk forthcoming a] Paulk, Mark C., Subrata Guha, William E. Hefley, Elaine Hyder, and Majid Iqbal. Forthcoming. *Comparing the eSCM-SP v2 and CMMI® v1.1: A Comparison Between the eSourcing Capability Model for Service Providers v2 and Capability Maturity Model® Integration v1.1*. Pittsburgh, PA: Carnegie Mellon University.

[Paulk forthcoming b] Paulk, Mark C., Subrata Guha, William E. Hefley, Elaine Hyder, and Majid Iqbal. Forthcoming. *Comparing the eSCM-SP v2 and Software CMM v1.1: A Comparison Between the eSourcing Capability Model for Service Providers v2 and the Capability Maturity Model® for Software*. Pittsburgh, PA: Carnegie Mellon University.

[PMBOK 2000] PMBOK. 2000. *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*. Newtown Square, PA: Project Management Institute.

[Prahalad 1990] Prahalad, C.K., and Gary Hamel. 1990. "The Core Competence of the Corporation." *Harvard Business Review,* May/June 1990.

[Pyzdek 2003] Pyzdek, Thomas. 2003. *The Six Sigma Handbook: Revised and Expanded*. New York, NY: McGraw Hill.

[Reifer 2001] Reifer, D.J. 2001. *Making the Software Business Case: Improvement by the Numbers*. Addison-Wesley.

[Scheaffer 1996] Scheaffer, Richard L., William Mendenhall III, and Lyman Ott. 1996. *Elementary Survey Sampling*. Fifth Edition. 2 Duxbury Press.

[Solingen 2004] Solingen, R.V. 2004. "Measuring the ROI of Software Process Improvement." *IEEE Software,* vol. 21, no. 3: 32-38.

[Stern 2001] Stern, J., and J. Shiely. 2001. *The EVA Challenge: Implementing Value-added Change in an Organization*. New York, NY: John Wiley & Sons, Inc.

[Stewart 1999] Stewart, Bennett G., III. 1999. *The Quest for Value*. HarperCollins.

[Wheeler 1992] Wheeler, Donald J., and David S. Chambers. 1992. *Understanding Statistical Process Control,* Second Edition. Knoxville, TN: SPC Press.

[Wheeler 2003] Wheeler, Donald J. 2003. *Making Sense of Data: SPC for the Service Sector.* Knoxville, TN: SPC Press.

Appendix A: An Overview of the eSCM-SP v2

This section provides a detailed overview of the eSourcing Capability Model for Service Providers (eSCM-SP) v2.

A.1. Rationale Behind Development of the eSCM-SP

IT-enabled sourcing, or eSourcing, uses information technology as a key component of service delivery or as an enabler for delivering services. It is often provided remotely, using telecommunication or data networks. These services currently range from routine and non-critical tasks that are resource intensive and operational in nature to strategic processes that directly impact revenues.

IT-enabled services are being sourced at a rapid rate. The evolution of the Internet and the global telecommunications infrastructure has provided client organizations with a choice of service providers located anywhere in the world. Simultaneously, competitive pressures have driven organizations to find the most cost-effective way to get the IT-enabled services they need while maintaining or improving their quality of service.

Sourcing failures are largely related to a core set of critical issues affecting sourcing relationships. Based on literature review [Kumar 2001] and interviews with eSourcing service providers and clients, issues critical for successful eSourcing have been identified. These include developing and sustaining stakeholder relationships, building and keeping a competent workforce, defining and delivering quality service, assessing and managing threats (e.g., disasters, invasion of networks), remaining competitive through innovation and improvement, and managing transitions of resources and services.

The combination of high growth and significant failures in eSourcing highlights a growing need: clients and service providers both need to be able to address the critical issues in sourcing in order to increase their probability of success. Individually and as a whole, existing frameworks do not address all of the critical issues in eSourcing. Also, many of these frameworks do not readily provide methods to assess the capabilities of IT-enabled service providers to establish, manage, and improve relationships with clients.

A.2. Structure of the eSCM-SP

Released in April 2004, the eSCM-SP v2 is composed of 84 Practices, which can be thought of as "best practices" associated with successful sourcing relationships. Each Practice is assigned a value along three dimensions: Sourcing Life-cycle, Capability Area, and Capability Level.

Each of the 84 Practices in the eSCM-SP contains information about a sourcing best practice. This information includes a statement summarizing the best practice, a description of the best practice, a list of activities needing to be performed, and supplemental information that helps clarify those activities. For more information on the structure of the 84 Practices, see [Hyder 2004a].

A.2.1. Sourcing Life-cycle

Although most quality models focus only on delivery capabilities, in eSourcing there are also critical issues associated with initiation and completion of an engagement. The first dimension of the eSCM-SP Practices highlights where in the Sourcing Life-cycle each Practice is most relevant. The Sourcing Life-cycle is divided into Ongoing, Initiation, Delivery, and Completion. Ongoing Practices span the entire Sourcing Life-cycle, while Initiation, Delivery, and Completion occur in specific phases of that Life-cycle.

Ongoing Practices represent management functions that need to be performed during the entire Sourcing Life-cycle. In order to meet the intent of these Practices, it is important to perform them across the whole life-cycle; an organization that only performs an Ongoing Practice during Delivery is not meeting the intent of the Practice. Initiation Practices focus on the capabilities needed to effectively prepare for service delivery. These Practices are concerned with gathering requirements, negotiating, contracting, and designing and deploying the service, including transferring the necessary resources. Delivery Practices focus on service delivery capabilities, including the ongoing management of service delivery, verification that commitments are being met, and management of the finances associated with the service provision. Completion Practices focus on the capabilities needed to effectively close down an engagement at the end of the Sourcing Life-cycle. They mainly include the transition of resources to the client, or to a third party, from the service provider.

A.2.2. Capability Areas

Delivery of eSourcing occurs through a series of interdependent functions that enables service providers to effectively deliver service. The second dimension of the eSCM-SP, Capability Areas, provides logical groupings of Practices to help users better remember and intellectually manage the content of the Model. These groupings allow service providers to build or demonstrate capabilities in each critical sourcing function, addressing all of the critical sourcing issues discussed above.

All of the Ongoing Practices are contained within six of the ten Capability Areas: Knowledge Management, People Management, Performance Management, Relationship Management, Technology Management, and Threat Management. The other four Capability Areas are temporal and are typically associated with a single phase of the Sourcing Life-cycle: Initiation, Delivery, or Completion. The exception is Service Transfer, which includes both Initiation and Completion Practices. In addition to Service Transfer, these temporal Capability Areas are Contracting, Service Design & Deployment, and Service Delivery.

The Knowledge Management Practices focus on managing information and knowledge systems so that personnel have easy access to the knowledge they need to effectively perform their work. This Capability Area addresses the critical issues of capturing and using knowledge, and measuring and analyzing reasons for termination.

The People Management Practices focus on managing and motivating personnel to effectively deliver services. They address understanding the organization's needs for personnel and skills, filling those needs, and encouraging the appropriate behaviors to effectively deliver service. This Capability Area addresses the critical issues of establishing

and maintaining an effective work environment, building and maintaining competencies, and managing employee satisfaction, motivation, and retention.

The Performance Management Practices focus on managing the organization's performance to ensure that the client's requirements are being met, that the organization is continually learning from its experience, and that the organization is continually improving across engagements. These Practices address the effective capture, analysis, and use of data, including data on the organization's capabilities relative to its competitors. This Capability Area primarily addresses the critical issues of maintaining competitive advantage, innovating, building flexibility, and increasing responsiveness. It also addresses monitoring and controlling activities to consistently meet service delivery commitments.

The Relationship Management Practices focus on actively managing relationships with stakeholders, including the client, as well as suppliers and partners who are integral to the delivery of services to the client. Relationship Management primarily addresses the critical issues of managing stakeholder expectations, establishing and maintaining trust and ensuring the effectiveness of interactions with stakeholders, managing supplier and partner relationships, managing the cultural differences between stakeholders, and monitoring and managing the client's and end-users' satisfaction. This Capability Area also addresses innovating, building flexibility, increasing responsiveness, establishing well-defined contracts with stakeholders, and maintaining a competitive advantage.

The Technology Management Practices focus on managing the availability and adequacy of the technology infrastructure used to support the delivery of the services. Their focus covers controlling the existing technology, managing changes to that technology, and appropriately integrating the technology infrastructure with the client, suppliers, and partners to effectively deliver service. This Capability Area addresses the critical issue of managing rapid technological shifts and maintaining technology availability, reliability, accessibility, and security. It also addresses innovating, building flexibility, and increasing responsiveness.

The Threat Management Practices focus on identifying and actively managing threats to the organization's ability to meet its objectives and the requirements of the client. They focus on active risk management, paying particular attention to the risks associated with security, confidentiality, infrastructure, and disasters that may disrupt service or fail to meet the requirements of the client. This Capability Area addresses the critical issues of managing clients' security, and ensuring compliance with statutory and regulatory requirements. It also addresses maintaining the continuity of service delivery, managing rapid technological shifts, and maintaining the availability, reliability, accessibility, and security of the technology.

The Contracting Practices focus on effectively managing the process of gathering client requirements, analyzing them, and negotiating a formal agreement that describes how the service provider will meet those requirements. A critical component of contracting is understanding the client's expectations and needs, and agreeing with the client on how the organization will meet those requirements. All Contracting Practices are in Initiation. This

Capability Area addresses the critical issues of translating implicit and explicit needs into the defined requirements, and establishing well-defined contracts with stakeholders.

The Service Design & Deployment Practices focus on translating the client's requirements and the contract language of what will be provided into a detailed design for how it will be provided, and on effectively deploying that design. This Capability Area is closely related to the Contracting Capability Area. All Service Design & Deployment Practices are in Initiation. This Capability Area addresses the critical issue of reviewing service design and deployment to ensure adequate coverage of the requirements. It also addresses developing procedures for monitoring and controlling activities to consistently meet service delivery commitments.

The Service Delivery Practices focus on the continued delivery of services according to commitments made to clients and based on service designs. They include planning and tracking of the service delivery activities. The Service Delivery Practices are the only ones in Delivery. This Capability Area addresses the critical issues of monitoring and controlling activities to consistently meet service delivery commitments, and maintaining continuity of service delivery. It also addresses establishing well-defined contracts with stakeholders, and maintaining a competitive advantage.

The Service Transfer Practices focus on transferring resources between service providers and clients or other service providers. In Initiation the resources are transferred to the organization as it takes responsibility for service delivery. This transfer may include people, processes, technology, and knowledge needed to effectively perform that service delivery. In Completion the organization transfers resources to the new service provider (either the client or an external service provider) in a manner that ensures continued service to the client during the transfer period. This Capability Area addresses the critical issues of smoothly transferring services and resources, and capturing and transferring the knowledge gained during the engagement to the client during contract completion. It also addresses maintaining continuity of service delivery.

A.2.3. Capability Levels

The third dimension in the eSCM-SP is Capability Levels. The five Capability Levels of the eSCM-SP describe an improvement path that clients should expect service providers to travel. This path starts from a desire to provide eSourcing services, and continues to the highest level, demonstrating an ability to sustain excellence.

The capabilities of Level 1 service providers vary widely. Some may have almost none of the eSCM-SP Practices implemented. These providers are very likely to be a high risk to work with because they often promise more than they deliver. Other service providers may have many of the eSCM-SP Practices implemented, including some Practices at Capability Levels 3 and 4. Because these service providers have not fully implemented all of the Capability Level 2 Practices, they may meet many of the client's needs successfully, but there will still be a risk of failure in areas where they have not implemented the necessary eSCM-SP Practices.

Service providers at Capability Level 2 have formalized procedures for capturing requirements and delivering the services according to commitments made to clients and other stakeholders. These providers are able to deliver specific services according to stated client expectations, given that the services do not significantly vary from the provider's experiences. At Capability Level 2 the service provider is able to systematically capture and understand requirements, design and deploy services to meet the requirements, and successfully deliver the services according to agreed upon service levels.

The infrastructure (e.g., work environment, training, technology, and information) is in place to support consistent performance of work that meets the service provider's commitments. Level 2 service providers have implemented all of the Capability Level 2 Practices and can demonstrate their effective usage.

Service providers at Capability Level 3 are able to deliver services according to stated requirements, even if the required services differ significantly from the providers' experience. At Level 3, the service provider is able to manage its performance across the organization, understand targeted market services and their varying requirements (including specific cultural attributes), identify and manage risks across engagements, and design and deliver services based on established procedures. The service provider supports this capability through sharing and using knowledge gained from previous engagements, objectively measuring and rewarding personnel performance, and monitoring and controlling technology infrastructure. Having established systems for forming and managing client relationships, providers at Capability Level 3 continuously aim to improve the services delivered. Improvements are reactive and are typically generated from the defined measurement and verification activities. The Level 3 service provider demonstrates measurable improvement with respect to organizational objectives. Organizational learning improves performance across engagements. Level 3 providers have effectively implemented all of the Level 2 and 3 Practices.

Service providers at Capability Level 4 are able to continuously innovate to add statistically and practically significant value to the services they provide to their clients and other stakeholders. At Capability Level 4 the service provider is able to customize its approach and service for clients and prospective clients, understand client perceptions, and predict its performance based on previous experiences. The service provider supports this capability through systematically evaluating and incorporating technology advances and setting performance goals from a comparative analysis of its current performance as well as from internal and external benchmarks. Level 4 providers systematically plan, implement, and control their own improvement, typically generating these plans from their own performance benchmarks. They have effectively implemented all of the Capability Level 2, 3, and 4 Practices.

Service providers at Capability Level 5 have demonstrated measurable, sustained, and consistent performance excellence and improvement by effectively implementing all of the Capability Level 2, 3, and 4 practices for two or more consecutive Certification Evaluations

covering a period of at least two years. There are no additional Practices required to reach Capability Level 5; effective, continued, implementation of all the eSCM-SP Practices in a rapidly changing environment shows an ability to sustain excellence throughout the organization over time.

A.3. Capability Determination Methods

ITsqc provides four methods that can be used to assess the capabilities of service providers relative to the eSCM-SP Capability Levels. The four Capability Determination Methods systematically analyze evidence of the provider's implementation of the eSCM-SP v2 Practices to determine what Capability Level their organization has achieved [Hyder 2004]. The Capability Determination may be of interest to, or required by, current or prospective clients of the service provider within a sourcing selection process. In this context, the Methods provide a consistent way for clients to evaluate their existing service providers or to compare two or more prospective providers. The knowledge from such an exercise based on eSCM-SP Capability Determination may be used to by clients to assess the risks and benefits of selecting a given service provider. Capability Determination may also be sponsored by service providers with the objective of evaluating their current capabilities and defining targets for self-improvement. In this context, the organization may or may not seek formal certification at an eSCM-SP Capability Level.

The four Capability Determination methods that are available from ITsqc are (1) Full Evaluation, (2) Full Self-appraisal, (3) Mini Evaluation, and (4) Mini Self-appraisal. The five major differences among these methods are (1) their purpose and outcome, (2) who does them, (3) who leads them, (4) who sponsors them, and (5) the number of eSCM-SP Practices that are analyzed (i.e., the model scope). Table 3 summarizes the four Methods.

Table 3
eSCM-SP Capability Determination Methods

		Evaluation	Self-Appraisal
Full	**Purpose**	For certification.	To prepare for a Full Evaluation or launch, or to validate an improvement effort. No certification.
	Team	External, trained and authorized by Carnegie Mellon University.	Internal, external, or combination.
	Lead Evaluator	Required.	Strongly Recommended.
	Sponsor	Client or service provider.	Service provider.
	Model Scope	All eSCM-SP Practices.	All eSCM-SP Practices.
Mini	**Purpose**	To prepare for a Full Evaluation, or as part of a provider-selection process. No certification.	To launch or validate an improvement effort. No certification.
	Team	External, trained & authorized by Carnegie Mellon University.	Internal, external, or combination.
	Lead Evaluator	Required.	Recommended.
	Sponsor	Client or service provider.	Service provider.
	Model Scope	Subset of eSCM-SP Practices.	Subset of eSCM-SP Practices.

Only the Full Evaluation leads to an ITsqc certification. It is a third-party external evaluation of a service provider's capability. It is based on evidence of the provider's implementation of all the Practices in the eSCM-SP, and is sponsored by the service provider or by its client(s). Members of the evaluation team must be trained by Carnegie Mellon University and must be authorized to perform external evaluations of service providers. An authorized Lead Evaluator must head the evaluation effort. The evaluation data are rigorously reviewed by a certification board at Carnegie Mellon University and, when warranted, results in certification by the university of the provider's capability. Organizations can be Certified eSCM-SP compliant at Capability Levels 2, 3, 4, or 5.

Appendix B: Considerations in Building the Measurement Report

B.1. Using the Goal-Question-Metric Paradigm to Define Measures

Two goals needed to be met to ensure the success of the eSCM-SP measurement effort. The first goal was to define example measures at the Capability Area level to support control and improvement within the context of the Model. This meant defining measures that most likely already existed within IT-enabled organizations, as well as identifying additional measures that were not currently being captured, but should have been. To define these measures the team looked to standard measurement information within the various measurement and quality standards. The second goal was to identify a number of value measures that an executive may use to measure whether a Practice or process (or the eSCM-SP itself) is adding value to the organization.

To meet these goals the measurement team used the Goal-Question-Metric (GQM) paradigm. The GQM is a top-down approach for the development, selection, and tailoring of measures. This paradigm ensures that there is a purpose for collecting each measure. Each measure provides valuable information—by answering a question—upon which decisions can be made and actions can be taken. By using GQM, the example measures contained within this report facilitate control and enable predictability to make decisions, take actions, make improvements, and provide consistency across an engagement within an organization—as well as ultimately across different market sectors.

NOTE: The sample Capability Area measures found within Appendix C do not focus on appraising the status of each Practice or Capability Area; this is accomplished through the Capability Determination process performed by assessors.

B.2. Role Perspectives

During the development of this eSCM-SP measurement document, the implementation of the GQM took on a subtly different approach than most implementations. To ensure that the "goals" would encompass all facets of a sourcing organization, the team needed first to define high-level sourcing roles. Although the role names may differ, these should be standard across the different market sectors within sourcing as well as standard to the five Levels of the eSCM-SP:

Organizational Manager

This role provides overall organizational management for the service provider. This role has the responsibility and authority to make decisions for multiple engagements within an organization.

Functional Manager

This role provides specific management designed to build and manage capability across all of the service provider's contracts. This includes Transition Management, Project Management, Account Management, Architecture, Knowledge Management, Telecom, Facilities, Marketing, Quality Assurance, Human Resources, and Security.

Engagement Manager
This role provides the interface to the client during the engagement.

Service Operations Teams
These are engagement- or project-specific service provision teams. They include the Service Delivery Team, Customer Service Representatives, Technical Support, etc.

B.3. Top-Down Approach: Value Added Measures

The top-level organizational goals defined for the Organizational Manager role translate a sourcing organization's expectations, strategic direction, and value proposition elements into a set of measures that focus on the top level of performance management. These top-down GQM measures allow management to monitor the overall performance and value being provided. They also provide insight into the value that a Capability Area, process, or the eSCM-SP is adding to the organization.

Why is it so important to have value measures in regards to the eSCM-SP? One of the primary drivers is so that an organization can effectively validate its sizable financial investment in the implementation of the eSCM-SP. Measuring the value-added gives an organization the ability to measure its changes in performance and financial gains that are linked to the implementation of the eSCM-SP.

B.4. Bottom-Up Approach: Capability Area Measures

The implementation of bottom-up Practice-level measures was another critical success factor for the implementation of the Model. By using the GQM, the top-level goals and measures cascade down to the process- and team-level roles (e.g., Functional Manager, Engagement Manager, and Service Operations Team) within the sourcing organization. Conversely, the goals identified for the team- and process-level roles align with and support the organization-level goals and role. These team- and process-level goals are translated into measures that are understandable and controllable at the team and individual level. Once these bottom-up measures are defined, the team reviewed the eSCM-SP Capability Areas and aligned them with the measures to determine whether they support the scope of work contained within the Capability Areas. These sample measures can be used by project teams not only to determine the effectiveness of the eSCM-SP Capability Areas, but also to measure a project or process, manage day-to-day activities, identify performance gaps, and drive process changes. When used together, the top-down value and bottom-up Capability Area measures should support implementation of the Model, true process change, and overall improvement. By using the GQM, this top-down and bottom-up approach ensures that the bottom-up measures are aligned with the overall organizational level goals.

B.5. The Balanced Scorecard

The Balanced Scorecard (BSC) defined by Kaplan and Norton [Kaplan 1996] is a tool that translates an organization's business objectives into a "balanced" set of performance measures. Using the Balanced Scorecard framework enables organizations to develop

an effective performance management system with measurements that look at service delivery, cost, and quality.

An effective Balanced Scorecard reflects the objectives of the organization by measuring performance from four key perspectives: financial, customer, internal business processes, and learning and growth. These different perspectives allow for a balance between short- and long-term objectives, and leading and lagging measures. They include different types of measures, such as service-level agreements, and process and product measures. A focus on balance is critical to the success of the BSC. To find this balance the organization needs to closely monitor not only the performance in each of the four perspectives—by itself, but also to determine how the performance in each of the perspectives affects the others.

For example, if employee satisfaction is low (low results within the learning and growth perspective), employees may not be motivated to provide adequate customer service (low results within the customer perspective), which in turn may affect the organization's financial results (low results within the financial perspective). Therefore, monitoring the performance in each of the perspectives and determining the cause and effect relationships between the measures within them reflects a full, balanced picture of what is occurring within the organization.

Listed below are descriptions for the four main perspectives that drive performance (as defined by Kaplan and Norton).

Financial Perspective
The measures within this perspective should indicate whether an organization's strategy is meeting its overall financial goals and objectives. To do this the measures should include revenue growth, productivity improvement and cost reduction, asset utilization, and financial risk management.

Customer Perspective
Within this perspective the customer's objectives and market-based strategy is translated into two sets of measures that reflect the factors that are most important to the customer. The first set typically includes core measures of market share, customer satisfaction and retention, profitability, and acquisition. The second set of measures should cover the attributes that an organization must deliver on to meet these core measures. They include service attributes of the delivered product (functionality, price, time, and quality), customer relationship attributes (response and delivery time), and image and reputation attributes.

Internal Business Processes
The development of measures for the internal business process perspective begins with identifying the end-to-end key internal business processes at which the organization must excel in order to meet the objectives of their customers. Once the processes have been identified, the organization must define measures that focus on the efficient definition, execution, and continuous improvement of these key processes. This should include process cost, quality, time, and performance measurements.

Learning and Growth

In the current business environment, targets for success keep changing. To stay successful, an organization must invest in its people, systems (including strategic information availability), and processes that build its capabilities. The measures in this perspective focus on how well the organization is meeting these objectives. They include the employee-focused measures of satisfaction, retention, productivity, reskilling and training, motivation, empowerment, roles and responsibilities, and alignment of the employee's goals with the organization's goals and objectives. To be effective, employees need to have excellent information so the organization must also measure its information system capabilities and technology infrastructure. Finally, the organization should include measures on team building and team performance. Due to the strong relationship between successfully meeting the organization's objectives within this perspective and attaining its goals in the other three perspectives, these learning and growth measures are critical to the overall success of the organization.

Measures, based on the strategic objectives of the organization, are defined for each of these perspectives. As mentioned previously, measures on the Balanced Scorecard should be an element in a chain of cause and effect relationships that will achieve the organization's strategic objectives. These relationships identify the tasks (causes) necessary to achieve business objectives (effect). Linkages within these chains are tested through the implementation of the BSC.

The BSC is not truly implemented until it is used to manage the organization. Using the Balanced Scorecard as a management tool results in an organization-wide understanding of objectives, strategies, and BSC measures and results, as well as regular reviews and discussions of the BSC by the organization's leaders and customers.

The development and implementation of a Balanced Scorecard can begin any time when implementing the eSCM-SP. However, by using it as early as possible, the organization is able to more effectively define measures within the different perspectives and manage the cause and effect relationships between these measures as the organization moves up the eSCM-SP Levels.

B.5.1. Scope of the Balanced Scorecard at the Different Levels of the eSCM-SP

The scope of the Balanced Scorecard depends upon where the engagement or organization is in its implementation of the eSCM-SP. At Level 2, the measures included within the BSC are engagement specific. However, as the focus of the eSCM-SP implementation moves from engagement specific (Level 2) to organization wide (Level 3 and Level 4), the measures should translate into an organization-wide BSC, to which the various engagement-specific BSCs will contribute. Although there are no Practices associated with Level 5, in order for the organization to achieve sustainable excellence it must continue to manage using its measurement program and be able to show improvement trends in achieving its business objectives as a result of implementing the eSCM-SP Practices. At this point the organization-wide Balanced Scorecard numbers should reflect improvement at sustained or improved levels.

B.5.2. The Relationship between eSCM-SP Practices, Measurement Categories, and Development of the Balanced Scorecard Perspectives

To facilitate the development of the BSC, including the four perspectives (financial, customer, internal business processes, and learning and growth), the organization should look to the various eSCM-SP Practices as well as the measurement categories defined in Section 4.5 (cost or effort, performance or satisfaction, status or progress, and nonconformance). The Practices and the measurement categories help with the identification of the measures within the various perspectives.

The following paragraphs include just a few examples of how the measurement categories can be included in each of the four perspectives. Please note that there is not a one-to-one relationship between the measurement categories and the four perspectives. The measurement categories can be included in one, two, three, or four of the perspectives. For example, performance or satisfaction measures can be included within the financial, customer, internal business processes, and learning and growth perspectives.

To develop the financial perspective the organization can consult the cost/effort category, which looks at the actual cost or effort that an activity requires. It can also refer to del08, "Financial management," which defines for the organization the financial controls that need to be in place to track costs for the organization. From this Practice an organization may define measures that track cost-benefit or earned value.

To develop the customer perspective the organization can refer to the measures within the performance/satisfaction category, which focuses on the effectiveness and efficiency with which work is being completed and the satisfaction of stakeholders with the results. Practice del03, "Deliver service," looks at using the service delivery plan to help ensure consistency in the delivery and achievement of committed service levels. From this Practice the organization may define measures for determining the percentage of service level agreements they meet. Another Practice, del04, "Verify service commitments," will also provide objective evidence to determine whether service-level agreements are being satisfied, and help identify any problems for which corrective action is required.

To develop the internal business processes perspective the organization can refer to the nonconformance category, which looks at conformance to policies, procedures, and processes at the required levels. Practice prf02, "Verify processes," looks at verifying that activities are being performed as expected and work products are complying with the requirements stated in procedures, standards, and other governance documents. From this Practice the organization may define measures that verify compliance to key business processes.

To develop the learning and growth perspective the organization can reference the status/progress category, which focuses on the status of activities, events, processes, and the progress that the organization is making on meeting their plans. Practice ppl08, "Personnel competencies," addresses employee skill gaps by defining training requirements and delivering the training. From this Practice the organization may define measures for tracking how well they meet their employees' training goals and plans.

B.5.3. eSCM-SP Practices and the Balanced Scorecard Perspectives

As noted previously, measures do not need to be defined for each and every eSCM-SP Practice. Instead, the organization needs to determine what is critical to measure in order to determine the success of meeting their customers' and the organization's business goals and objectives, and to help manage the organization's business and eSCM-SP implementation. Once the organization determines what is important to measure, and which Practices are relevant, in many cases the detailed Activities within these Practices (and ultimately the measures that the organization defines to cover these Practices) may fit into one, two, three, or all four of the perspectives.

For example, tfr02, "Personnel transferred in," could have measures in three of the perspectives. They could include employee satisfaction (measuring an employee's satisfaction level with the transfer process), as well as measuring unplanned turnover. Both of these measures could be included within the learning and growth perspective. Another measure tied to this practice could be customer satisfaction, measuring the customer's satisfaction with their personnel's transfer into the organization. This measure could be included within the customer satisfaction perspective. Finally, a plan versus actual measure of the personnel transfer plan and number of issues identified during the execution of the transition of personnel process could be included within the internal business processes perspective. Please refer to the detailed descriptions within Chapters Three, Four, Five, and Six in *The Balanced Scorecard: Translating Strategy Into Action* [Kaplan 1996] for additional information on categorizing measures within the four perspectives and for example Balanced Scorecards.

B.6. Goal-Driven Measurement Applied

While preparing this report, once roles were defined, measures were defined that measure at the Capability Area level, measure value, reuse industry standards, and work to effectively manage a sourcing organization. First, goals were defined based on the organization roles. Accordingly, questions were asked to help determine whether an organization was fulfilling those goals. These questions were answered by defining specific appropriate measures. Finally, the Capability Areas within the Model were then aligned with the defined measures. This approach produced the following results:

- Capability Areas were aligned with the overall goals of a sourcing organization.
- Measures were defined that align to a number of the Capability Areas.
- The number of example measures that this report suggests was greatly reduced.

To better understand the approach taken, please refer to the goals defined for each role, shown in Table 4.

Table 4
Goals for Each Role

Role	Goals
Organizational Manager	Maximize profitability
	Provide growth
	Reduce costs
	Increase referenceability
	Improve predictability
	Reduce risks
Functional Manager	Provide required resources
	Maintain security, confidentiality, and intellectual property compliance
	Maintain statutory and regulatory compliance
	Introduce innovation
	Meet infrastructure availability needs
	Provide an adequate working environment
	Ensure work products, processes and service commitments are met
	Build workforce competencies
	Design/develop/deliver service according to specifications
	Meet commitments
	Optimize performance of technology infrastructure
	Smooth transition of resources
	Manage risks
	Minimize impact of disaster
	Meet or exceed organizational expectations
Engagement Manager	Meet client requirements
	Meet budget targets
	Add customer value (business objectives)
	Maintain and grow customer relationship
	Ensure team performance
	Manage engagement scope
	Manage risks
	Minimize impact of disaster
Service Operations Team	Meet or exceed customer expectations
	Meet or exceed organizational expectations

Appendix C: Operational Definition Template and Sample Measurements

An operational definition is a clear, concise, detailed explanation of a measurement. Operational definitions are fundamental to data collection, particularly if a decision is being made on the quality of an outcome and there is room for confusion. It is easy to assume that people collecting data understand what and how to complete their tasks. However, people have different views and opinions and these affect the data collection. The only way to ensure consistent data collection is by means of detailed operational definitions that eliminate ambiguity. When collecting data, it is essential that everyone in the system has the same understanding and collects data in the same way. Operational definitions should therefore be specified before collection of data begins.

The following template demonstrates the use of an operational definition in the eSCM-SP:

Measure Name		
Type (Cost/Effort, Status/Progress, Nonconformance, Performance/Satisfaction)		
Purpose of Measure	**Description of Use:**	Defines the use of this measure.
	Goal:	The goal(s) that needs to be achieved, based on the goal-driven measurement approach.
	Questions:	Question(s) that help determine whether or not the goals have been met.
Measurement Definition:	**Formula:**	A clear, concise, and detailed explanation of the measure can include any of the measures to the left.
	Unit of Measure:	
	Frequency:	
	Exclusions:	
Decision Criteria:		Describes acceptable versus unacceptable behaviors of the measure. Provides guidance on interpreting the data.
Data Collection:		Describes indicative methods for data collection: survey, audit, etc.
Additional Comments:		States any additional information that will aid in explaining this measure.

The examples below demonstrate the use of the operational definition template. Each corresponds to the example by the same number in Appendix D.

Example 1: Daily Average Speed to Answer (ASA)		
Performance/Satisfaction Measure		
Purpose of Measure	**Description of Use:**	Measures the average time that callers waited on hold during the measurement period.
	Goal:	Provide service at a given level.
	Questions:	Does the service provided meet the contractual requirements? Is the service consistent within a range?
Measurement Definition:	**Formula:**	$\text{ASA} = \frac{\text{Total Wait Time of All Calls}}{\text{Total Number of All Calls}}$
	Unit of Measure:	ASA: time per call Number of Calls: count Wait Time: seconds
	Frequency:	Daily
Decision Criteria:		1) If the value of the measure exceeds the contractual level on any given day, then the target is deemed to have been missed. (Voice of the Customer) 2) Any point outside the control limits requires investigation to determine the cause for the change. (Voice of the Process)
Data Collection:		ASA is computed automatically and captured daily by the telephone switch. The data are collected daily via ftp and saved in the data warehouse. All data are displayed in SPC format.
Additional Comments:		None.

Example 2: Mean Time Between Outages (MTBO)		
Performance/Satisfaction Measure		
Purpose of Measure	**Description of Use:**	Measures the mean (average) time between outages across all systems.
	Goal:	Improve system stability.
	Questions:	How long, on average, do systems stay up between failures?
Measurement Definition:	**Formula:**	$\text{MTBO} = \frac{\text{Availability} \bullet \text{MTTR}}{\text{(1-Availability)}}$ $\text{Availability} = \frac{\text{Time System is Operational}}{\text{Total Time in Interval}}$ $\text{MTTR} = \frac{\text{Total Time to Repair}}{\text{Total Number of Repairs}}$
	Unit of Measure:	MTBO: days Availability: none (ratio of times) MTTR: mean time to repair (hours)
	Frequency:	Monthly
	Exclusions:	1) Systems designated as research are not included in this calculation. 2) Scheduled down time is not included in the Total Time in Interval or the Time System is Operational.
Decision Criteria:		Any run of 8 consecutive points all above the center line indicates a positive change in the process target. (Voice of the Process)
Data Collection:		Availability and MTTR are captured daily from problem management logs. MTBO is calculated monthly after all input data have been entered.
Additional Comments:		This is an internal operational measure.

Example 3: Backup Failure Rate

Nonconformance Measure

Purpose of Measure	**Description of Use:**	Measures the fraction of backups that fail.
	Goal:	Improve backup reliability.
	Questions:	What is the rate of backup failures?
Measurement Definition:	**Formula:**	$\textbf{Failure Rate} = \frac{\textbf{Backups Not Completed Successfully}}{\textbf{Total Backups Scheduled}} \bullet 100$ (See Additional Comments)
	Unit of Measure:	Failure Rate: percentage Backups: count
	Frequency:	Daily, by date of the start of the backup window.
Decision Criteria:		Any run of 8 consecutive points all below the center line indicates a positive change in the process target. (Voice of the Process)
Data Collection:		Backups scheduled are recorded in the daily operational scheduling system. Backups completed successfully are recorded in the daily operational logs. Backups scheduled with no corresponding success record within the backup window are considered failures. The backup window spans midnight, so the reporting date for each backup window is the date the window starts.
Additional Comments:		1) This is an internal operational measure. 2) A backup must complete within the nightly backup window to count as successful.

Example 4: Customer Satisfaction

Performance/Satisfaction Measure

Purpose of Measure	**Description of Use:**	Measures the percent of customers who are satisfied with the service.
	Goals:	1) Increase customer satisfaction. 2) Meet contractual service level.
	Questions:	How satisfied are customers with the service?
Measurement Definition:	**Formula:**	$\textbf{Percent Satisfied} = \frac{\textbf{Satisfied Responses}}{\textbf{Total Responses}} \bullet 100 \pm CI$ $CI = 1.96 \bullet \sqrt{\frac{p(1-p)}{n}}$, *where* p= percent satisfied n=number of surveys entered
	Unit of Measure:	Percent Satisfied: percentage Responses: count
	Frequency:	Weekly
	Exclusions:	Surveys where the respondent chooses "No Opinion" are omitted from the analysis.
Decision Criteria:		1) Any week where the entire confidence interval for the percent satisfied falls below the target is considered a failure to meet the contractual service level. 2) Any run of 8 consecutive points all above the center line indicates a positive change in the process target. (Voice of the Process) 3) Any point outside the control limits requires investigation to determine the cause for the change. (Voice of the Process)
Data Collection:		Customer satisfactions surveys are stored in the data warehouse as they arrive. Weeks run from Sunday through Saturday. Satisfied responses are those responses marked either "Very Satisfied" or "Satisfied."
Additional Comments:		1) This is a contractual measure. 2) The confidence interval (CI) is a 95% confidence interval, hence the constant of 1.96 in the formula above.

Example 5: Security Breach Rate		
Status/Progress Measure		
Purpose of Measure	**Description of Use:**	Measures the security breach rate each time a breach occurs.
	Goal:	Understand security breach exposure.
	Questions:	Are breaches becoming more frequent?
Measurement Definition:	**Formula:**	$\text{Breach Rate} = \frac{365}{\text{Days Since Last Breach}}$
	Unit of Measure:	Breach Rate: breaches per year
	Frequency:	Irregular
Decision Criteria:		1) Any point outside the limits indicates a significant change in the breach rate. 2) Any run of 8 consecutive points all above or all below the center line indicates a sustained change in the breach target. (Voice of the Process)
Data Collection:		Breaches are recorded by date in the security database.
Additional Comments:		1) This is an internal operational measure. 2) Measure is calculated at each breach occurrence.

Example 6: Effort to Perform Application Management		
Cost/ Effort Measure		
Purpose of Measure	**Description of Use:**	Measures the effort required to perform application management.
	Goal:	Understand staffing requirements for application management service.
	Questions:	How many staff-hours are required to maintain an application of size x?
Measurement Definition:	**Formula:**	**Staff = Size/750**
	Unit of Measure:	Staff: number of maintenance personnel required Size: function points counted using IFPUG standard
	Frequency:	Irregular
Decision Criteria:		1) If available skills are less than Staff, then hiring may be necessary.
Data Collection:		Actual effort spent on the engagement are tracked in a data repository to update the formula as necessary. Actual size is measured on a quarterly basis to understand the application growth over time.
Additional Comments:		1) This is an internal operational measure. 2) Model is updated at regular intervals based on historical data.

Figure 6
XmR Chart for Average Speed to Answer

Appendix D: Examples

D.1. Example #1: Call Center Measures, Service Level Indicators, and Operational Measures

Call centers typically perform a lot of measurements. Modern telephone technology has allowed them to capture almost every aspect of their business that relates to the telephone interaction itself. Consequently, finding data in this service area is seldom a problem, but turning that data into useful information is in many cases still a challenge.

In many call centers, primary attention is focused on measures of service quality such as Average Speed to Answer (ASA), Percent Answered Within an Interval, Abandon Rate, and First Call Resolution. These measures, or similar ones, should be tracked at regular intervals with sufficient granularity to be able to detect significant shifts as quickly as possible. Because these are regularly-collected time series data, they lend themselves naturally to Statistical Process Control (SPC) charts.[1] In order to determine that a change is significant, the measurement methodology must separate signals from the normal noise inherent in any data set, and SPC does just that. Figure 6 shows an Individuals and Moving Range (XmR) control chart of Average Speed to Answer over a 30-day interval.

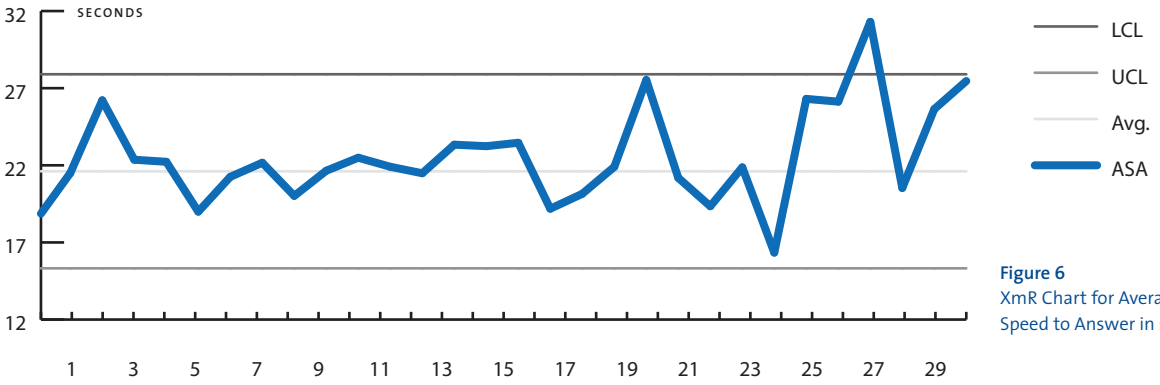

Figure 6
XmR Chart for Average Speed to Answer in seconds

The chart shows that this measure averages 21.6 seconds and varies between 15.3 seconds and 27.9 seconds daily, except day 27. To determine why day 27 is outside the upper limit, what was different about that day, additional Operational Measures may come into play. By itself, this Service Level Indicator chart does not tell us why day 27 was different; it only tells us that we need to focus our efforts there. Thus, Operational Measures become critical to actually understanding and controlling the business. As important as the Service Level Indicators are—they are the customer's requirements in most cases—they are insufficient by themselves.

There are three primary factors that affect the ASA: the volume of calls arriving to the center, the average time spent by agents handling those calls, and the number of agents staffed to handle the calls. Other subtle factors can and do influence ASA, but these are the

[1] The formulas used for calculating control limits and performing other forms of statistical analysis are not specified in this report. A number of books are available on these topics for those interested in specific analyses [Wheeler 1992, Wheeler 2003].

significant ones. Therefore, efforts should be focused on measures of those primary factors to further narrow the search. For the sake of brevity, we assume that the call volume chart in Figure 7 was the only chart that showed an exception on day 27.

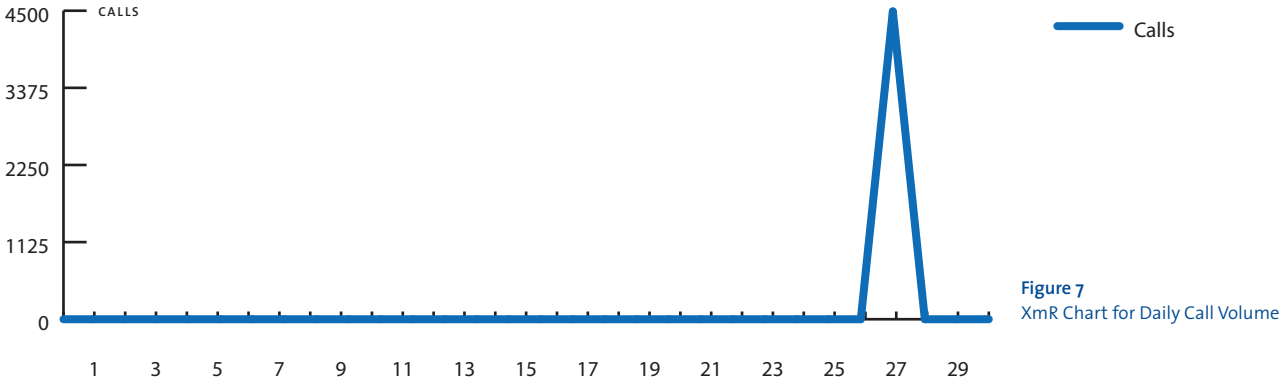

Figure 7
XmR Chart for Daily Call Volume

Figure 7 leads to the conclusion that the excess volume on day 27 was the primary cause of the higher ASA on that same day shown in Figure 6. Had other charts also shown exceptions that day, there would have been additional factors to investigate. While this example was quite straightforward, it clearly demonstrates the need to have Operational Measures to provide the needed insight to understand changes in Service Level Indicators.

D.2. Example #2: Mainframe Measures Using Relationships to Generate Measures

Measures related to availability and reliability of the system(s) are among those most frequently collected and reported in a mainframe environment. Three such measures are Availability (AVAIL), Mean Time to Repair (MTTR), and Mean Time Between Outages (MTBO).

The most common measure, AVAIL, is calculated by dividing the time a system is up-and-running in a given time interval by the total time in that interval. As such, it is always a value between 0.0 and 1.0 (it is often converted to a percentage by multiplying by 100), and AVAIL can be computed for any time interval desired.

MTTR, on the other hand, can only be computed in intervals where at least one outage has occurred. This is because MTTR is the ratio of total outage time (the total time to repair) over the total number of outages; if no outages occurred in an interval, then MTTR is undefined. For single systems, this is often an issue, but when computing MTTR for large groups of systems, there is often at least one outage in a reasonable interval, such as a month. Thus, in most environments, AVAIL and MTTR can be measured and plotted in SPC form just as ASA was in Example #1.

The third measure, MTBO, creates a special set of problems. If we attempt to compute MTBO in the same manner that we computed MTTR (i.e., by using the total time between outages and dividing by the number of outages), we encounter a dilemma in both the

single-system and the multiple-system cases. In the single-system case we cannot compute MTBO if there have been no outages. Fortunately, in this situation AVAIL contains all the information that is useful. In the multiple-system situation, though, because it is less likely that all systems are 100% AVAIL, the question of how to treat systems with no outages arises. If only the systems with outages are used in the MTBO computation, the measurement is biased to the low side, because systems that were always up had no impact on the measure. Adding the 100% AVAIL systems into the calculation is not easy, however. For any given interval, their last outage and their next outage fall outside the current interval, thus the amount of time one would need to add is impossible to know. So, how does one obtain MTBO?

It turns out that these three measures are related mathematically, and this relationship can be exploited to generate MTBO from AVAIL and MTTR:

$$\text{AVAIL} = \frac{\text{MTBO}}{\text{MTBO+MTTR}}$$

Therefore, solving for MTBO, we have the following result:

$$\text{MTBO} = \frac{\text{AVAIL} \bullet \text{MTTR}}{(1 - \text{AVAIL})}$$

Clearly, if MTTR is undefined, then this computation is not usable either, but that means that AVAIL must be 100% or 1.0 and the system(s) are flawless in the interval. What was a difficult measure to compute directly now becomes simple to compute using the other two measures.

It should be noted that an MTBO that has been calculated is really an "instantaneous" MTBO (i.e., it applies only over the interval for which it was calculated). As such, it is possible that this MTBO will be somewhat volatile from period to period—a characteristic that would be quickly discerned through the use of control charts.

D.3. Example #3: Server Backups Dealing with Counts and Rates

In today's complex server environments, performing backups is a significant and time-consuming task, and every organization needs to have an effective backup process in place. Further, organizations need to gather data to verify conformance to the backup process and understand its effectiveness. One measure that is often reported is the number of backup failures each day (i.e., the difference between the planned number of backups and number of successful backups). In the worst case, a failed backup represents a data exposure until the next successful backup occurs, and could result in the inability to recover data. In the best case where failed backups are re-run in a timely manner, the exposure is lessened, but the failed backup still generates additional cost because of the rework required to correct the situation. In order to understand such an environment and to make improvements over time, it is necessary to be able to characterize the failures.

For this example, the data in Table 5 were collected over a 30-day period for a moderately large server environment:

Table 5
Failure Data by Day

Day	1	2	3	4	5	6	7	8	9	10
Failures	5	3	2	2	7	7	8	7	6	3
Day	11	12	13	14	15	16	17	18	19	20
Failures	7	6	4	4	6	4	7	6	7	9
Day	21	22	23	24	25	26	27	28	29	30
Failures	2	6	9	5	6	9	2	5	3	11

Day 30 shows the highest number of failures during the interval, and the question arises, "Is this exceptional?" In order to answer this question, one must determine the normal variation for these values using a control chart. The average count per day is well above 1.0, and we assume that the number of backups attempted is reasonably consistent each day. Therefore it is appropriate to place these counts on an XmR chart. The result is shown in Figure 8.

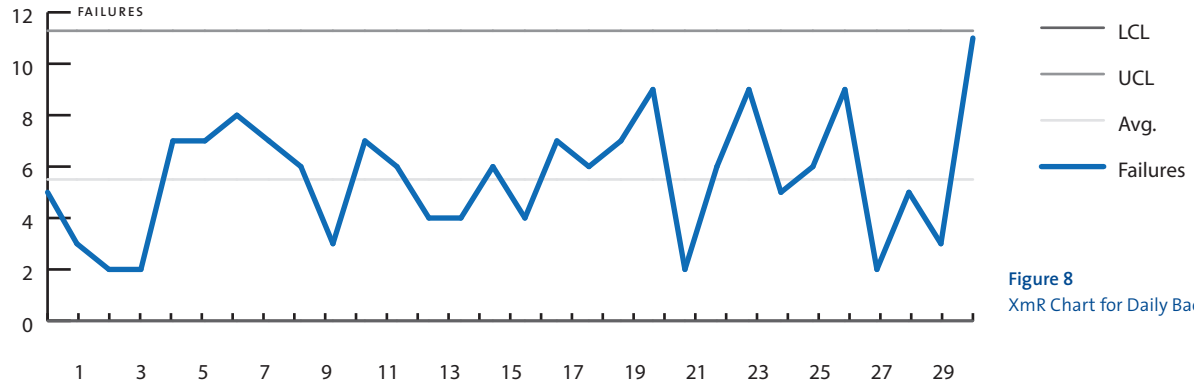

Figure 8
XmR Chart for Daily Backup Failures

In this case the system averages 5.5 failures per day, and the number of failures could be as high as 11.3 due to normal variability as shown by the upper limit. Thus, the number of failures on day 30 is not excessive, given the current state of the backup process. Since the chart shows no exceptional variation, work can now begin on trying to improve this process as a whole with the goal of reducing the number of failures in the future.

What if the number of backups performed each day had not been fairly consistent? In a complex environment it is not uncommon to have backup schedules that change depending on the day of the week. If that is the situation, then it is inappropriate to chart the failures directly, because each day's failures relates to a differing number of backup opportunities. In that case, one must convert the counts to rates by dividing each day's count by the number of attempts on that day. These rates may then be plotted on XmR charts. Suppose the backup counts in Table 6 applied to the data in Table 5.

Table 6
Backup Opportunities by Day

Day	1	2	3	4	5	6	7	8	9	10
Attempts	7,559	6,551	5,884	6,606	7,101	5,444	8,065	6,394	5,226	7,722
Day	11	12	13	14	15	16	17	18	19	20
Attempts	5,248	7,116	6,355	6,535	6,800	5,993	7,750	5,729	7,696	6,320
Day	21	22	23	24	25	26	27	28	29	30
Attempts	7,573	6,555	4,044	7,279	5,873	7,024	6,235	5,865	5,536	6,516

Turning the counts and attempts into rates yields the XmR chart in Figure 9. Now day 23 demonstrates an excessive failure rate and demands investigation. By using rates instead of counts, we are able to compare data even when the areas of opportunity are not similar. A good rule of thumb is to use rates whenever the largest count is more than 1.5 times the smallest count. When in doubt, it is always safe to convert to rates, even if the areas of opportunity are identical.

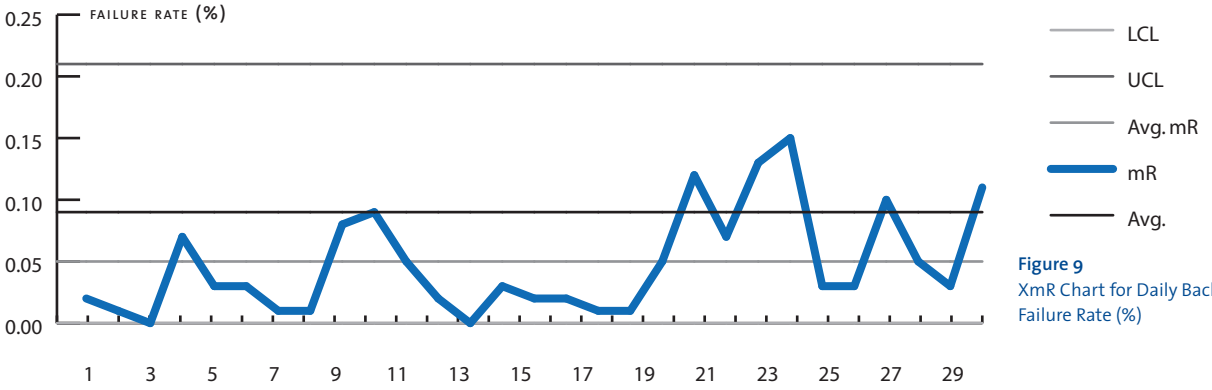

Figure 9
XmR Chart for Daily Backup Failure Rate (%)

D.4. Example #4: Customer Satisfaction Measures the Quality of the Measure

Using surveys to determine customer satisfaction is a standard part of many service businesses. The results of these surveys are put to a variety of uses, and the numerical results are often disseminated widely within a company, among its customers, and even to the public at large. While many surveys are reported along with a statement that the results are accurate to "plus or minus x percent," many include no such disclaimer. This inconsistency leads to at least two questions. First, just what does such a statement mean, and second, is such a statement really necessary?

To answer the first question it is necessary to examine what a sampling survey is. In the case of a customer satisfaction survey there is a group of people, the customers, whose opinions the organization wishes to understand. In an ideal world, every customer would be asked how he or she feels about the service, and the results would be tallied. While this sounds appealing on the surface, it is impractical for a number of reasons. Issues such as cost, time, and customer unwillingness to be surveyed, make it virtually impossible to survey every customer. As a result, a technique known as sampling is used to try to understand the views of customers.

2 Bias introduced by non-response can be a significant problem in surveys, but it will not be discussed here.

3 The concept that a statistic such as the mean of repeated samples varies less than the original data is a result of the Central Limit Theorem.

In sampling a random subset of the overall customer population is examined to infer characteristics of the population as a whole. Typically, some percentage of the overall population is randomly selected to receive a survey. Of those selected, some fraction actually responds. These responses are then analyzed to produce the results of the survey. The percentage of customers that are satisfied or the average score are commonly determined as a measurement of interest. The values are computed, and can vary depending on the people selected to complete the survey and the ones who choose to respond.[2] For example, if a company has 1,000 customers, a 20% sample was surveyed, and 50% of those surveyed responded, then the response group size would be 100 people (1,000 x 0.2 x 0.5). Clearly, there are many different samples of size 100 that are possible with a group of 1,000 persons. Many of these samples will give different results for a measure such as percent satisfied. It is these different results that generate the need for confidence intervals (i.e., intervals around the estimate that are likely to contain the actual value we are seeking).

The statistic that is computed from a sample is an estimate of the true value of the parameter we seek for the entire population. If every member of the population were sampled and all responded, the estimate would be exact, and the confidence interval would have zero width. In other words, the exact value of the parameter would be known. When sampling, however, the estimate contains uncertainty precisely because only a subset of the population is considered. It is this uncertainty that the confidence interval captures. When an estimate has, for example, a 95% confidence interval of plus or minus 2%, it means that, if the process were repeated many times, the true value of the parameter would lie between the estimate plus 2% and the estimate minus 2% approximately 95% of the time.

The amount of the uncertainty in the estimate varies according to three things: the amount of variation in the parameter being estimated, the size of the sample, and the size of the population itself. If the population parameter varies considerably with respect to the characteristic we are examining, then the estimates generated by our sampling can vary significantly as well, though less than the population variance.[3] Conversely, if the population varies little, so do the estimates.

The size of the sample affects the amount of uncertainty as well. In general, the smaller the sample, the larger the confidence interval (i.e., the less precise the estimate is). This is also intuitive, as smaller samples are more susceptible to being unrepresentative of the population as a whole.

Finally, the size of the population has an effect because of its interaction with the sample size. As the percentage of the population sampled increases, the confidence interval decreases. The relationship among these three factors is well understood and is defined by the formula below, which represents the approximate 95% confidence interval of an estimate of the mean for a sample of size *n* from a population of size *N* [Scheaffer 1996].

$$CI = \pm 2\sqrt{\frac{s^2}{n}\left(\frac{N-n}{n}\right)}$$

Here, s^2 is the sample variance of the parameter whose mean we are estimating. With the formula it is clear that the size of the confidence interval is dependent on the three variables: s^2, n, and N. While N is known ahead of time, the values of s^2 and n cannot be determined until the responses have been gathered and tabulated. Typically, a desired confidence interval is chosen before the survey and an estimate of s^2 is used with the formula above to determine the sample size, n, that will give the desired accuracy. In spite of this estimation, until the actual data are available, the final confidence interval cannot be determined. The sample may not turn out to have the response that was anticipated, or the variance may not be what was estimated, thus changing the final confidence interval.

When attempting to compare an estimate to a fixed value such as a Service Level Target, it is necessary to account for the width of the sample-specific confidence interval. If we have a target, such as 90%, and an estimate of 88.7% with a plus or minus 2% confidence interval, then we cannot say that the estimate is distinguishable from the target of 90% because the interval includes the target. In other words, the estimate is not sufficiently accurate to distinguish it from 90% in this example. Assuming the variance stays constant, future estimates can be made more accurate by increasing the sample size, thereby attempting to look at a larger portion of the population. There is, in most cases, a cost associated with increasing the sample size. By reporting and using confidence intervals correctly, it is clear to all who use the data that the values are truly estimates and, as such, they have a degree of variability that must be considered in decision-making. To do otherwise is misleading.

D.5. Example #5: Security—Dealing with Infrequent Events

One measure that is often tracked but not widely disseminated is the number of security breaches of a given system.[4] While breaches are not desirable, they occur occasionally in many environments. A common way of reporting these events is seen in Figure 10.

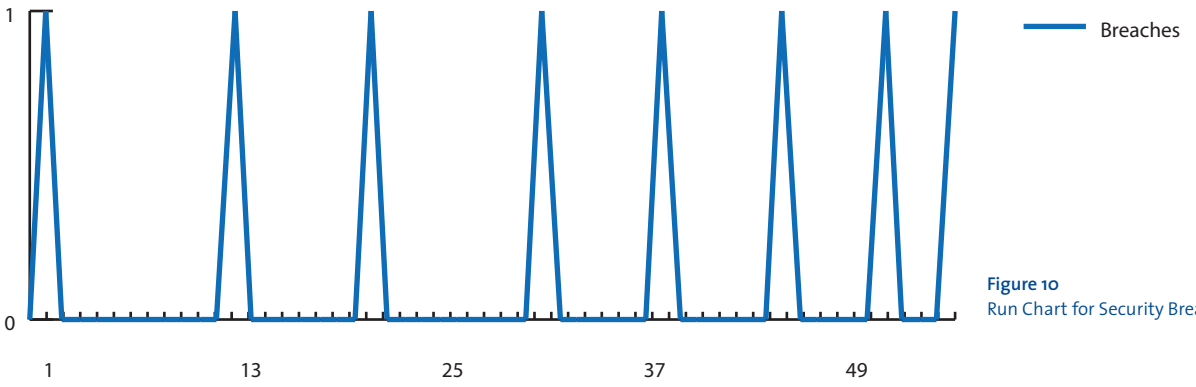

Figure 10
Run Chart for Security Breaches

While this chart shows that there have been 8 breaches in the last 55 months, it doesn't clearly help us to understand whether the problem is getting better, worse, or staying the same over time. The average number of breaches per month is 0.125, so months where even a single breach has occurred are significantly worse than the average. This is hardly

4 Adapted from an example in *Making Sense of Data* by Donald J. Wheeler, SPC Press, 2003, pp. 229-232.

insightful. While it is possible to find SPC charts that can be used on this data set directly (XmR charts do not work with it at all), no SPC chart type works very well because the average frequency is below 1.0 events per interval. An alternate approach must be taken.

Instead of looking at the number of breaches, we can look at their rate of occurrence. This can be done by counting the number of days between each of the breaches and taking the reciprocal to arrive at a breach per day rate. The dates of the breaches are shown in Table 7.

Table 7
Dates of Breaches

Date of Breach	2/23/01	1/11/02	9/15/02	7/6/03	2/19/04	9/29/04	3/20/05	7/13/05
Day of Year	54	11	258	188	50	272	79	194
Days Between Breaches		322	247	295	227	222	172	115

These data are converted to rates in Table 8. The rates are rescaled to breaches per year to make them more manageable, but this does not affect the conclusions. The first five intervals have been chosen to compute the limits on the XmR chart for breach rate.

Table 8
Breach Rate

Days Between Breaches	322	247	295	227	222	172	115
Breaches per Day	0.0031	0.0040	0.0034	0.0044	0.0045	0.0058	0.0087
Breaches per Year	1.13	1.48	1.24	1.61	1.64	2.12	3.17

The XmR chart in Figure 11 shows the result of the analysis. Because relatively few data points were used to compute the control limits (5 points), these limits are soft. In spite of this, the second breach in 2005 is well outside the upper control limit and is a signal that the breach rate has increased. This chart gives much more insight into the nature of the problem than did Figure 10. It should be noted that this methodology assumes that the breaches come from a single cause system. If the systems generating the breaches are known to be different over time, than the whole concept of common-cause variation is nonsensical, and the chart will provide no insight into how to improve the system.

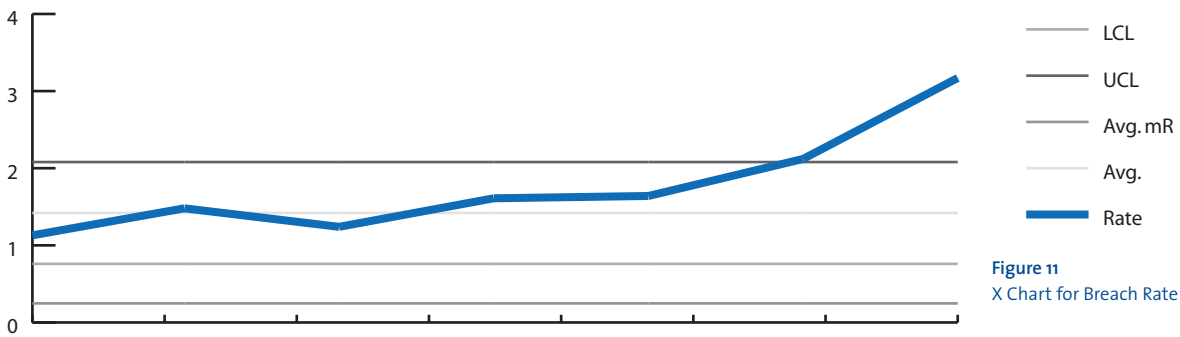

Figure 11
X Chart for Breach Rate

D.6. Example #6: Planning—Dealing with Changes to the Agreement

Figure 12 illustrates the types of measurements that may be used to support an IT-enabled sourcing engagement. In its simplest form, the engagement can be viewed as a "black box" that accepts a volume of work requests and produces a volume output of responses. The volume of work determines the size of the engagement (e.g., calls for a call center, function points for an application management, jobs for a mainframe). The length of time needed to complete the work is the responsiveness. The work is produced for an overall cost, and productivity is calculated as the work output divided by cost. (The ability of the volume output to pass its acceptance standards is a measure of its quality.) A nonconformance is an output that fails to meet its acceptance criteria. Each of these factors should be a part of the organizational performance baseline.

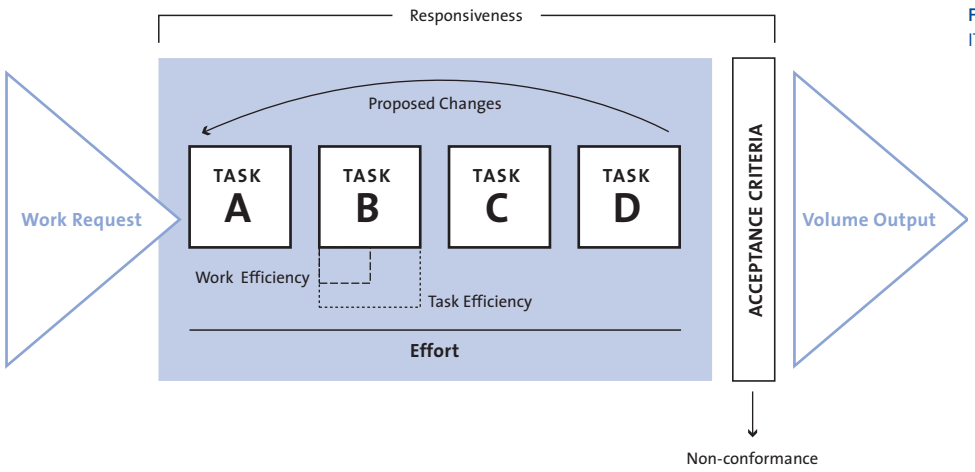

Figure 12
IT-enabled Measurement Context

The service provider is able to control factors like process steps, task and work efficiency, internal rework levels, staffing and costs. By manipulating these items, a service provider influences its costs, capacity, responsiveness, and productivity. Quality issues related to the output from the service may create a backlog of work.

Other factors that determine the volume of work requests are solely under control of the client. These include official requests following standard processes and "under the table" work which passes directly between members of the client and service provider teams. Identifying and quantifying the "under the table" work is difficult, yet is an important challenge for the organizational performance process. Since this work is not officially sanctioned, it is invisible to client managers and it fails to be included in plans. This failure is often the root of later client dissatisfaction with the service provider.

The standard processes for addressing official requests cover analysis, design, implementation, and deployment of software upgrades and defect corrections on the client's IT infrastructure. It is common for the client to ask the service provider to provide estimates for the cost, schedule, and quality of a proposed project, including management of an additional application to the contract.

The service provider measures the size of the application using an agreed-to standard—for example, the international function point user group (IFPUG) standard for counting function points. The outcome of this count is then entered into an empirically derived formula to arrive at the number of staff required to support the new application. As requirements are added, deleted, and refined for an application, the service provider must manage these work requests and measure their size, effort, efficiency, cost, and responsiveness to improve their future estimates and document the effects of process improvements.

Appendix E: Measurement from the Perspective of Capability Areas

This section lists key measurement issues and objectives to be considered by adoptees of the eSCM-SP as they develop their own set of suitable measures. A Capability Area-based measurement perspective may be useful for organizations that use a structure similar or analogous to the eSCM-SP Capability Areas to assign management roles and responsibilities, or to define improvement objectives. These matrices allow organizations to sort and organize Practices by categories of measure, wherever that might provide a useful perspective for measurement, control, or improvement. Measures may not be necessary for every Practice in a Capability Area. Conversely, a given Practice may require more than one measure for adequate control.

The key measurement issues and objectives are organized by Capability Area, across four categories of measures:

- Cost/effort (C/E)
- Status/progress (S/P)
- Nonconformance (NC)
- Performance/satisfaction (P/S)

For each Practice listed, a "1" in a category of measures column indicates that the category is of primary concern. A "2" indicates secondary concern. It must be noted that these rankings are neither absolute nor definitive. Their purpose is to provide a quick (at-a-glance) view of which category of measures is primarily at play within a given Capability Area. A numerical scheme (x.y) is used, where x refers to the number of the Practice, and y refers to one of the four categories of measure. The comments refer to the corresponding (x.y) issues or objective.

Knowledge Management (knw)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
knw01: Share knowledge				
knw02: Provide required information	2	2	2	1
knw03: Knowledge system	2	1		2
knw04: Process assets	2	1		2
knw05: Engagement knowledge		1		2
knw06: Reuse				1
knw07: Version & change control	1	2	2	
knw08: Resource consumption			1	2

Issues and objectives

(1.1) Progress or status on (i) provision of information required for delivery and support activities, (ii) identification of gaps in required information, and (iii) remedial actions from previous reviews.

(1.2) Satisfaction ratings from users on (i) knowledge system and (ii) process assets.

(2.1) Problems or issues attributed to (i) unauthorized changes or versions, (ii) ineffective controls, or (iii) ineffective procedures.

(2.2) Cost/effort estimates per unit per category of change.

(3.1) Variance in resource consumption by (i) resource type, (ii) contract or client, and (iii) service type.

Comments

(1.2) Quality attributes should include (i) accuracy, (ii) completeness, (iii) timeliness, (iv) currency, and (v) security.

(1.2) Users include process owners and service personnel.

(2.1) Failures in service delivery or support attributable to unauthorized or unsuccessful changes may be counted here.

People Management (ppl)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
ppl01: Encourage innovation				
ppl02: Participation in decisions				
ppl03: Work environment	1			2
ppl04: Assign responsibilities		1		2
ppl05: Define roles		1		2
ppl06: Workforce competencies		1		2
ppl07: Plan & deliver training		1	2	2
ppl08: Personnel competencies		1		
ppl09: Performance feedback		1		
ppl10: Career development				1
ppl11: Rewards				1

Issues and objectives

(1.1) NCs related to statutes or regulations with respect to work environment.

(1.2) Employee satisfaction with work environment and support services.

(2.1) Progress or status on determining (i) competency needs, (ii) competency gaps, and (iii) plans for eliminating gaps or meeting needs.

(2.2) Problems or issues attributed to (i) assignment of roles and responsibilities and (ii) completeness or effectiveness of required training.

(3.1) Variance in training costs and schedules from budgets and plans .

(3.2) Satisfaction ratings on effectiveness of training from (i) trainees or (ii) super ordinates or sponsors.

Comments

(1.2) Qualitative measures and indices may be required for measuring employee satisfaction.

(2.1) Need to develop Personnel Competency Inventory (PCI) and define what constitutes a "competency gap."

(3.1) There should be some cost and effort basis for training that would vary by (i) type or mode, (ii) location, or (ii) subject matter of training.

Performance Management (prf)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
prf01: Engagement objectives	2	1		
prf02: Verify processes	1	2		2
prf03: Adequate resources		1	2	2
prf04: Organizational objectives	2	1		
prf05: Review organizational performance		2		1
prf06: Make improvements		2	2	1
prf07: Achieve organizational objectives		2		1
prf08: Capability baselines		2		1
prf09: Benchmark		2		1
prf10: Prevent potential problems		2	2	1
prf11: Deploy innovations		1	2	

Issues and objectives

(1.1) Process commitments breached within control perspective (period, function, phase, etc.).

(1.2) Problems or issues with performance of processes attributed to (i) NC to procedures, (ii) poor design or definition of processes, or (iii) delays and shortfalls in required resources.

(2.1) Difference between (i) capability baselines and actual performance and (ii) expected and actual performance and satisfaction ratings received from clients.

(2.2) Patterns and trends in organizational performance across periods, contracts, regions, etc.

(3.1) Progress on corrective or preventive actions on problems or issues related to (i) performance, (ii) resources, and (iii) baselines.

Comments

(1.1) OLAs or internal SLAs may be in place.

(1.2) Shortfalls include delays.

(2.1) Should highlight (i) under-performance or need to revise baselines and (ii) problems with requirements and expectations.

(2.2) Consolidated and multi-D views may reveal chronic and systemic problems .

(3.1) Focus on effectiveness and efficiency of problem management process.

Relationship Management (rel)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
rel01: Client interactions	2			1
rel02: Select suppliers & partners	1			
rel03: Manage suppliers & partners	2		2	1
rel04: Cultural fit				1
rel05: Stakeholder information		1		2
rel06: Client relationships			2	1
rel07: Supplier & partner relationships			2	1
rel08: Value creation		2		1

Issues and objectives

(1.1) Problems or issues related to interactions with clients or suppliers attributed to (i) NC of policies or procedures, (ii) cultural fit, or (iii) quality of information available on stakeholders.

(2.1) SLAs breached by suppliers (by period, contract, customer, etc.).

(2.2) NCs in procedures for evaluation or selection of suppliers.

(3.1) Satisfaction ratings from stakeholders (including clients and suppliers) on categories such as (i) quality of interactions, (ii) responsiveness, (iii) problem resolution, and (iv) cultural fit.

(3.2) Cost/effort for unit measure of performance and satisfaction.

Comments

(2.1) SLAs with clients typically backed by SLAs or OLAs with suppliers and internal service organizations.

(3.1) Regular surveys and combination of qualitative and quantitative data on satisfaction, performance, and perceptions.

Technology Management (tch)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
tch01: Acquire technology	2	1	2	2
tch02: Technology licenses	1			2
tch03: Control technology	1	2	2	
tch04: Technology integration		1	2	2
tch05: Optimize technology		2	2	1
tch06: Proactively introduce technology		1	2	2

Issues and objectives

(1.1) Problems or issues attributed to shortfall or constraints in (i) capacity, (ii) performance or functionality, (iii) unauthorized changes, (iv) unsuccessful integration, or (v) other complications.

(2.1) Underutilized or overloaded capacity not attributable to client-side change in requirements or demand characteristics.

(3.1) Change (+/-) in unit costs attributable to changes, upgrades, consolidation, or optimization of technology.

Comments

(1.1) Problems or issues: SLAs breached, delays in deployment or delivery, unplanned costs, vulnerabilities, risks, reduced efficiencies, etc.

(3.1) May be based on service delivery units or transactions.

Threat Management (thr)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
thr01: Risk management				
thr02: Engagement risk	2	1		
thr03: Risk across engagements	2	1		
thr04: Security	1	2		
thr05: Intellectual property	1	2		
thr06: Statutory & regulatory compliance	1	2		
thr07: Disaster recovery	1	2		

Issues and objectives

(1.1) Risk exposure assessed by control perspective.

(1.2) Change (+/-) in risk exposure from last review attributed to (i) mitigation, (ii) change in regulations, (iii) change in technology or infrastructure, or (iv) scope of contract or service definition.

(2.1) Number of NCs and trends by (i) severity and (ii) type, or (iii) control perspective.

(2.2) Cost of actual NCs or exposures suffered by (i) severity and (ii) type, or (iii) control perspective.

(3.1) Problems or failures flagged on disaster recovery plans based on testing or actual occurrences.

Comments

(1.1) Based on (i) assets or commitments, (ii) threats, and (iii) probabilities and impacts.

(1.1) Control perspectives: contract, client, location, regulation, etc.

(2.1) Type: external, internal, operational, regulatory, IP, etc.

(2.2) Cost types: penalties, remedial, productivity, damages, ratings, etc.

Contracting (cnt)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
cnt01: Negotiations	2			1
cnt02: Pricing	1			2
cnt03: Confirm existing conditions	2	1		2
cnt04: Market information		2		1
cnt05: Plan negotiations		1		2
cnt06: Gather requirements	2	1	2	2
cnt07: Review requirements	1	2	2	2
cnt08: Respond to the requirements		2	2	1
cnt09: Contract roles		1		
cnt10: Create contracts		1	2	
cnt11: Amend contracts				

Issues and objectives

(1.1) Quality of negotiated positions.

(2.1) Problems or issues attributed to (i) NCs with pricing policies or guidelines, (ii) quality of cost estimates, models or algorithms, or (iii) ineffective estimation or review processes.

(2.2) Unplanned costs and unaccounted risks due to problems or issues in (1.1) above.

(3.1) Client satisfaction with respect to quality responses to requests.

(3.2) Design, deployment, or delivery issues attributed to (i) contract development or (ii) requirements management processes.

(4.1) Variance in cost/effort for governance of client and supplier contracts.

Comments

(1.1) Covers contracts and agreements. Quality of positions should consider Pareto efficiency.

(3.1) Such as RFPs, RFIs, and RFCs.

(4.1) May include business development costs, legal and administrative overheads, and cost of governance activities.

Service Design & Deployment (sdd)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
sdd01: Communicate requirements		1		2
sdd02: Design & deploy services		1	2	2
sdd03: Plan design & deployment		1	2	
sdd04: Service specification		1		2
sdd05: Service design	2	1	2	2
sdd06: Design feedback		2		1
sdd07: Verify design	2	2		1
sdd08: Deploy service		1	2	2

Issues and objectives

(1.1) Problems or issues attributed to (i) poor design, (ii) NC to design guidelines or procedures, (iii) ineffective or missing design reviews, or (iv) poor planning or execution.

(2.1) Progress and performance on design and deployment plans.

(2.2) Unplanned costs or delays due to problems as defined in (1.2) above.

(3.1) Client satisfaction with deployment process and results (e.g. cost and schedules).

Comments

(1.1) Identified from (i) design reviews and feedback, (ii) project reviews during deployment, and (iii) analysis of problems during ongoing delivery and support.

(2.1) Project management methods and measures may be applied such as CPM, EV, and cost and schedule variances.

Service Delivery (del)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
del01: Plan service delivery		1		
del02: Train clients		1	2	2
del03: Deliver service	2	2	2	1
del04: Verify service commitments	2			1
del05: Correct problems		2	2	1
del06: Prevent known problems		2	2	1
del07: Service modifications	2	1	2	2
del08: Financial management	2		1	2

Issues and objectives

(1.1) Progress and performance on (i) service delivery planning, (ii) service modifications, and (iii) contract amendments.

(2.1) SLAs met within control perspective.

(2.2) Problems or issues attributed to (i) NC to procedures, (ii) availability of resources, (iii) changes implemented, (iv) ineffective monitoring, or (v) ineffective or missing service level reviews.

(2.3) Unplanned costs or delays due to problems in (2.1) and (2.2).

(3.1) Critical problems by (i) severity of impact, (ii) escalation levels or age, and (iii) category of service or client, and their associated impact.

(4.1) Client satisfaction with delivered service.

(5.1) Variance in (i) unrecovered or unallocated costs, (ii) unauthorized expenditure, and (iii) actual and estimated costs.

(5.2) Service problems or issues attributed to (i) shortfalls in budgets or (ii) delays in funding or authorization.

Comments

(1.1) Control perspective: period, client, region, etc.

(2.2) Other factors may include (i) change in demand characteristics, (ii) effects of major disruptions, and (iii) supply or capacity problems.

(3.1) Criteria for criticality need to be defined. Factors include SLA commitments and financial implications.

(4.1) Dependent partly on the number of SLA performances.

Service Transfer (tfr)

eSCM-SP v2 Practice	NC	S/P	C/E	P/S
tfr01: Resources transferred in	2	1	2	2
tfr02: Personnel transferred in		1	2	2
tfr03: Service continuity		2		1
tfr04: Resources transferred out	2	1	2	2
tfr05: Personnel transferred out		1	2	2
tfr06: Knowledge transferred out	2	1	2	2

Issues and objectives

(1.1) Progress and performance on transfer of resources and control.

(1.2) Problems or issues with transfers related to (i) unaccounted or misappropriated resources, (ii) transfer of accountability and control, (iii) intellectual property ownership and use, or (iv) poor planning or estimation of cost, and effort required for transfers.

(2.1) Cost and schedule variances for transfer projects.

Comments

(1.1) Project management methods and measures may be applied.

5 All of these case studies ignore the more complex issues of accounting for the value of capital expenditures and assume the cost of an item is its cost.

Appendix F: Examples of Measuring Value

The following three examples demonstrate how organizations at different Capability Levels analyze the value added to the organization through implementation of eSCM-SP Practices.

F.1. A Level 2 Example—Implementing the Threat Management Practices

For the organization aspiring to Capability Level 2, decisions are most likely made at the level of the individual engagement. Consequently, in many cases, value decisions are made at the same level. In this example, a single-engagement organization is considering implementing the Level 2 Threat Management Practices of the eSCM-SP. Threat management is critical to success in the sourcing arena, and six of the seven Practices in this Capability Area are at Level 2. Only thr03, "Risk across engagements," is not a Level 2 Practice. Thus, the organization is faced with the decision of determining the value for implementing the six Level-2 Threat Management Practices. While it had bits and pieces of each of the Practices, the organization needed to make an investment in order to fully implement all of the Level 2 Threat Management Practices.

The organization decided to use a Net Present Value (NPV) approach to determine which projects to undertake. Further, after studying the Threat Management Practices, it decided to combine them into three groups (projects) for the NPV analysis:

- Group 1: thr01, thr02 – Project 1
- Group 2: thr04, thr05, thr06 – Project 2
- Group 3: thr07 – Project 3

This was done for two reasons. First, the organization felt that the Practice groupings were logically related, allowing them to be more efficient in the way they approached implementation. Second, the organization felt that it could not easily separate the benefits from the Practices within each grouping, so the task of determining benefits was made easier by the groupings. This is a common way to deal with the problem of confounding in cases where the parts can be logically considered as part of some larger whole.

To perform an NPV analysis, the organization needed to turn all of the costs and benefits into dollar amounts. The costs were easy to turn into dollars, as most of the cost for each of the projects was either people's time or equipment that needed to be purchased.[5] The benefits, on the other hand, were more difficult to quantify. The team assigned to do the cost-benefit analysis determined that benefits flowed to the projects in three forms: reduction in current effort on the existing engagement, reduction in penalties paid on the existing engagement, and increased opportunities to sell services to new customers. The reduced effort numbers were estimated by reviewing the Practices and determining where having a consistent process would likely reduce such items as errors and rework in the current environment. Those time estimates were in hours, so they were easily converted to dollar amounts. The penalty risks were easy to determine, based on an analysis of historical payments that had been made. The expected value of future quarterly payments was computed and used as a benefit where appropriate. Finally, the team determined that by

documenting its disaster recovery procedures (thro1 – group 3), the company would be able to compete for and win business they had previously not been qualified for. Therefore, a conservative estimate for new revenue was entered as a benefit for Project 3. The longest any group would need to have funding was four quarters, so the team decided to look at five quarters for their NPV analysis. The costs and benefits by quarter for each project are shown in Table 9.

Table 9
NPV Costs and Benefits by Quarter for Each Project

	Project 1			Project 2			Project 3		
	Cost	**Benefit**	**Net**	**Cost**	**Benefit**	**Net**	**Cost**	**Benefit**	**Net**
Q1	$10,000	$0	($10,000)	$8,000	$0	($8,000)	$5,000	$0	($5,000)
Q2	$2,000	$500	($1,500)	$8,000	$0	($8,000)	$5,000	$0	($5,000)
Q3	$0	$5,000	$5,000	$8,000	$10,000	$2,000	$30,000	$0	($30,000)
Q4	$0	$10,000	$10,000	$6,000	$10,000	$4,000	$0	$20,000	$20,000
Q5	$0	$15,000	$15,000	$0	$10,000	$10,000	$0	$25,000	$25,000
NPV	$11,594	$26,871	$15,277	$27,960	$26,662	($1,297)	$37,022	$39,335	$2,313

The NPV analysis used a 3% per quarter cost of capital. Looking at the NPV values in the last row and the column labeled "Net" for each of the projects shows a mixed story. Clearly, Project 1 shows significant value over the five-quarter period with a NPV of $15,277. While not as attractive, Project 3 also has a positive NPV. Because the organization used conservative revenue projections for the analysis of Project 3, it felt comfortable that this was a "worst-case" five-quarter return, so this project was also approved. Project 2, however, had a negative NPV in the five quarters that were considered. In further discussions, the organization decided that the $10,000 per quarter benefit would continue for the foreseeable future for this project, so it too would have a positive NPV after only one more quarter. Had the team computed the payback period for each project—the period required for the project's benefits to exceed the project's costs—they would have seen that Project 2 became profitable in Q6. In light of Project 2 having a positive NPV of $7,077 after only one more quarter, the organization decided to fund all three projects.

F.2. A Level 3 Example—Creating a Measurement Repository

For the organization aspiring to Capability Level 3, the decision-making focus must move up to the organizational level from the individual engagement level. This is not to imply that all decisions are outside the scope of any single engagement, but rather that decisions about policies, procedures, and investments need to be coordinated across engagements so that consistency is achieved across the organization as much as possible. Consequently, value decisions need to be made at a higher level as well. In this example, a multi-engagement organization is considering implementing the Level 3 Knowledge Management Practice, knw04, "Process assets." While each engagement in the organization has its own measurement repository, there is no centralized repository for data. As a result, organization-wide analyses have been rarely performed, and when they were, they

were time-consuming undertakings. This is one of the main areas where the organization feels that it is not meeting Level 3 standards, and it decided to perform a value analysis to determine if it should move to a single repository.

Unfortunately, funding is scarce, and the marketing department has also requested a similar amount to fund a new advertising campaign. The organization can not afford both investments, so it is important to determine which one provides the better value. This organization decided to use the NPV approach described above to make its investment decision between these two competing opportunities.

The organization has seven active engagements, ranging from one that had been signed just two months ago and was still in transition, to their first engagement that they had successfully serviced for over five years. With the exception of the newest engagement, measurement systems on the engagements are adequate for the task of delivering the immediate service requirements. Each engagement is independent, though there has been some sharing of tools, spreadsheets, and database queries across some of the more recent engagements.

The details of the advertising investment are not important here. Using a 2.5% per year discount rate, the NPV analysis yields an 8.3 million dollar NPV over the three-year life of the advertising campaign—not a bad number in its own right. This is the bar that the measurement repository needs to exceed.

Because the NPV method requires the costs and benefits for the projects, the organization needs to consider these two aspects for the measurement repository. While the benefits are easy to list, they are more difficult to quantify. Determining the time savings for each engagement in the areas of collecting, scrubbing, and analyzing data is fairly straightforward. Other benefits are less easily determined. For instance, how much value do you place on being able to compare data across engagements? What is increased customer loyalty worth? How much is it worth to know that some parts of the organization are less efficient than others? While these are difficult questions, the organization needs to find ways to quantify the benefits as much as possible. The organization uses an estimate of the benefits based on expected cost reductions it feels it can achieve as a result of increased efficiencies gained from identifying and acting on opportunities for improvement. While these numbers are clearly estimates, the organization uses a standard set of questions and assumptions to generate the estimates, to maintain as much rigor as possible in the process. The organization looked at the time remaining on the contracts for each engagement to determine the length of time to use in the analysis.

The costs, on the other hand, were reasonably straightforward. The organization knew what it would take to perform a detailed design of the repository, build it, and move each of the engagements into it. It was determined that the repository could be operational within 15 months, and all groups could be converted within two years. The net cash flows for each engagement and the organizational overhead are shown in Table 10.

Table 10
Net Cash Flows for the Organization and Each Engagement

Cash Flows (thousands)	Year 1	Year 2	Year 3
Engagement A	($3,000)	($5,000)	($5,000)
Engagement B	($3,000)	$8,000	$10,000
Engagement C	($3,000)	$6,000	$9,000
Engagement D	($3,000)	$6,000	$9,000
Engagement E	($3,000)	$2,000	$7,000
Engagement F	($3,000)	$2,500	$4,000
Engagement G	($3,000)	$10,000	$5,000
Overhead	($25,000)	($7,000)	($2,000)
Total	($48,000)	$24,500	$37,000

A problem arose as soon as the costs and benefits were quantified. The group involved with Engagement A, the oldest and most prestigious engagement in the organization, decided that this project made no sense for them and they did not want to participate. Each engagement had targets to meet, and these costs jeopardized the Engagement A group's attainment of those targets. As the table indicates, the cash flow for all three years is negative for Engagement A; they would get little benefit as their measurement system was quite mature, and the development and conversion would provide them with negative cash flow every year. On the other hand, if the total value was better than the advertising project, the repository was clearly valuable to the organization as a whole. In fact, using the 2.5% discount rate and the cash flows from Table 10, the NPV of the repository project was projected to be 10.9 million dollars, making it the preferred project. The organization was faced with a common challenge for an aspiring Level 3 organization: globally optimal decisions are often locally sub-optimal for some groups. In this case, every engagement group except A would be better off as a result of the conversion to the measurement repository, and the organization as a whole would be much better off as well. Not having the benefit of the Engagement A group's data would have been a serious shortcoming for the repository effort, as that group's processes were the most mature in the company.

In the end, the organization modified their incentive system and made the right decision to move all the engagements to the new repository. The internal incentive plans had been a barrier to making the right decision, and the organization realized this and corrected it. Had the decision been made at the engagement level with each group having the ability to opt out, the result would have been an inferior one for the organization.

F.3. A Level 4 Example—Deciding Among Alternative Investments

For the organization aspiring to Capability Level 4, the decision-making focus remains at the organizational level, but it becomes more sophisticated. At this point the organization has a mature data collection and analysis capability, and it is expected to continually use those assets to enhance their performance and increase their value to their customers. Because of these factors, investment decisions need to account for factors that were often ignored (both consciously and unconsciously) by organizations at lower Capability Levels. The most notable of these factors are the concepts of risk and uncertainty:

The term risk refers to a situation in which the potential outcomes can be described in objectively known probability distributions. Risk is a measure of the probability and severity of adverse effect. The term uncertainty refers to a situation in which no reasonable probabilities can be assigned to the potential outcomes. Uncertainty is the inability to determine the true state of affairs of the system. [Haimes 1998, pg. 228]

In this example, an organization has an opportunity to make investments in several projects that will enable it to create value for its customers (rel08, "Value creation," is a Level 4 Practice). Each of the organization's projects has a variety of possible financial outcomes, ranging from profits to losses. Further, the projects are not independent of each other, and that interaction affects the possible outcomes.

From a financial and operational standpoint, the organization determined that the four potential projects—A, B, C, and D—could be undertaken in only the following combinations:

- A & B
- A & C
- C & D

The organization begins by doing an NPV analysis on each of these project pairs. While its analysis considers more factors and deals with many more engagements than the one performed by the aspiring Level 2 organization above, it is nevertheless a similar approach. Unfortunately, the results are not conclusive. Using expected values for all of the inputs, the NPV results are all in the $15 million range, plus or minus less than 1 million dollars. Given the risks in each of the project pairs, the organization is uncomfortable making a decision when the results are so close.

It is in situations like this that the opportunities of the higher Level organizations manifest themselves. Because this organization has significant historical data on previous investments of these types, it is able to embark on a more sophisticated analysis of the problem. The organization uses its data to build a Monte Carlo simulation of the three investment options it is considering. A Monte Carlo simulation is a technique that allows an organization to replace static assumptions with probability distributions, thereby developing a more complete model that captures the risks involved in decisions. Further, the organization is able to capture the relationships between the projects in each pair, and more accurately model those interaction effects as well. While the simulation gives the organization many new insights into its decision, Figure 13 below shows the two most significant results.

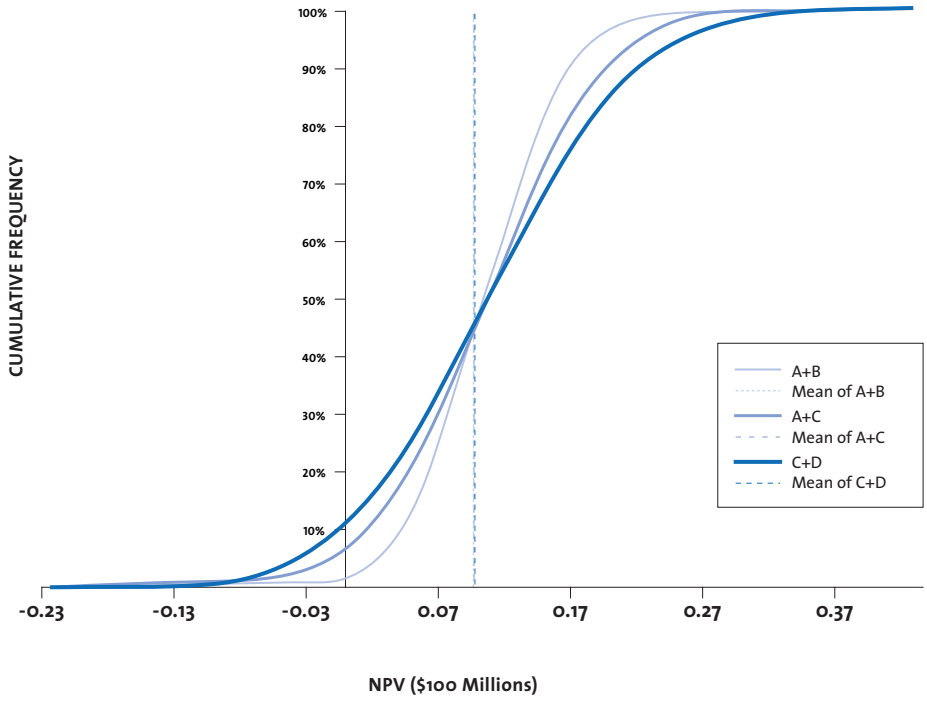

Figure 13
Cumulative Probability of NPV for the Three Project Pairs

Figure 13 shows the cumulative probability of various returns for each of the three project pairs. The first result is that the simulation shows the average NPV to be only $10 million, not the $15 million the static analysis had shown. Because the model captures the nonlinear nature of the problem, the average input levels did not lead to average outputs. Further, even though the average values are almost identical, the risk levels of the three project groups are quite different. C & D is the most risky pair, having a much larger range of possible NPV outcomes. A & B is the least risky pair, with A & C being in the middle. Now the organization has the insight they need to choose among the three options that seemed indistinguishable. While Monte Carlo techniques are clearly available to all of the organizations in these three examples, it is the availability of high-quality historical data that enable this soon-to-be Level 4 organization to take full advantage of these techniques.

Appendix G: Measurement in Related Standards and Models

The following models and standards are briefly discussed from a measurement perspective: ISO 9001, COBIT, BS 15000 and ITIL, CMMI [Chrissis 2003], Software CMM [Paulk 1995], and COPC. Detailed discussions of the relationships between the eSCM-SP and these and other standards and models are available or under development [Paulk forthcoming, Guha 2005, Guha forthcoming, Iqbal forthcoming, Iqbal 2004, Paulk forthcoming a, Paulk forthcoming b]. One goal of the measurement effort was to use "standard" definitions and terminology whenever possible.

G.1. Measurement in ISO 9001 (Quality Management Systems—Requirements)

ISO 9001 promotes the adoption of Deming's "Plan-Do-Check-Act (PDCA)" approach for process improvement [Deming 1986, ISO 2000]. The "check" element of PDCA requires continuous monitoring and measurement of process and product quality.

All requirements stated in this international standard are generic and can be applied to any type of business that involves a contractual agreement with customers. ISO 9004:2000 provides additional guidance to help organizations implement the requirements of the ISO 9001 standard for performance improvement [ISO 2000a].

Requirements for measurements are stated in clause 8 (Measurement, analysis, and improvement) of ISO 9001. The purpose of measurements as stated in the standard is to demonstrate product conformity, process conformity, and continuous improvement of the quality management system.

The standard expects organizations to collect three categories of measures: customer satisfaction (clause 8.2.1), process quality (clause 8.2.3), and product quality (clause 8.2.4). The word "product" also includes service. The standard does not prescribe specific measures or methods. It is left to the organizations to identify the measures suitable for their business. It is expected that the identified measures should be adequate to demonstrate whether the processes and products are conforming to the requirements. The standard is very rigid about product quality. All products must conform to acceptance criteria before release. Organizations have to establish a defined methodology for measurement, analysis, and corrective actions. Process and product compliances are verified through audits (internal and external).

ISO 9004 provides some additional guidance on measurement. Possible sources of customer satisfaction data are addressed in clause 8.2.1.2. Converting process and product measurement data to financial information to enable ROI calculation is performed in clause 8.2.1.4. Clause 8.2.2 provides additional guidance on process measurement including some example attributes for measuring process performance (e.g., cycle time, throughput). These performance measures should be used to evaluate process performance and initiate corrective actions when performance deviates from expected result. Clause 8.2.3 provides detailed guidance on product quality measurement. Product quality measurement should be performed according to a predetermined plan. ISO 9004's section 8.2 also provides some

example methods for product quality verification (e.g., type testing, in-process inspection). Use of statistical techniques is strongly recommended for data analysis (clause 8.1.2, 8.2.1.3, 8.4).

G.2. Measurement in COBIT

The Control Objectives for Information and related Technology (COBIT) is a framework for governance, control, and audit for the information and related technology considered necessary and useful for IT organizations to effectively meet the needs and obligations of their stakeholders. COBIT was developed and reviewed by a global committee of experts, organized by the IT Governance Institute (ITGI), who achieved consensus on a set of "good practices" that would help organizations put in place a control system or framework to manage IT processes to effectively support business processes [ITGI 2000]. COBIT identifies 34 IT processes across four domains, a high-level approach to achieve control over each process, 318 detailed control objectives, and audit guidelines to evaluate the performance of the processes. Organizations can use the guidance that COBIT provides to define, implement, and monitor the appropriate level of control within their IT organizations. Each of the 34 IT processes belongs to one of the following domains:

- Planning and Organization (PO)
- Acquisition and Implementation (AI)
- Delivery and Support (DS)
- Monitoring (MO) Audit Guidelines

For each IT process and high-level control objective, COBIT specifies one or more of the following business requirements for information that must be satisfied to effectively support business needs: effectiveness, efficiency, confidentiality, integrity, availability, compliance, and reliability. COBIT aims to provide guidelines for management to answer the following questions about their organizations [ibid.]:

- What are good indicators of performance?
- What is important from a control perspective?
- What are the critical success factors for control?
- What are the risks of not achieving our objectives?
- What do others do? How do we measure and compare?

COBIT recommends the use of the Balanced Scorecard approach to frame goals of the business that are to be achieved with the support of information technology enablers. The outcomes of business processes are represented on Balanced Business Scorecards as Key Goal Indicators (KGI) that inform management—after the fact—whether an IT process has achieved its business requirements, usually expressed in terms of the following information criteria [ibid.]:

- availability of information needed to support the business needs
- absence of integrity and confidentiality risks

- cost-efficiency of process and operations
- confirmation of reliability, effectiveness, and compliance

The performance of IT processes in enabling business processes to achieve the desired outcomes is measured on the IT Balanced Scorecard in the form of Key Performance Indicators (KPI), which provide a basis for judging whether or not the business goals will be reached. KPIs should be defined so that they provide a reliable indication of performance that is to be controlled for IT processes, to effectively support the attainment of goals of business processes as measured by KGIs. For each of the 34 IT Processes, COBIT provides KPIs and KGIs for management to consider.

Since the primary concern of IT processes is to deliver information required by business processes on a timely basis, management needs to identify and control the factors critical to success in achieving control over performances and outcomes [ibid.]. For each for the 34 IT Processes, COBIT provides a set of Critical Success Factors (CSF) that define the most important issues or actions for management to achieve control over and within its IT processes. These serve as guidelines for implementation and a check-list for management to consider from multiple perspectives of managing their organizations.

With respect to achieving control over IT process, COBIT recommends benchmarking processes against their high-level control objectives to determine the following [ibid.]:

- current status of the organization—where it is today
- current status of (best-in-class) the industry—the comparison
- current status of international standard guidelines—additional comparison
- organization's strategy for improvemen—where it wants to be

COBIT recommends that management use a maturity model based on Carnegie Mellon University's Software Engineering Institute (SEI℠) to assess the level of control in a given IT process and to compare it with levels achieved by other organizations or the industry. COBIT recommends that each of the 34 IT Processes be rated on a scale of 0 to 5, ranging from "nonexistent" to "optimized".

G.3. Measurement in BS 15000 and ITIL

The BS 15000 standard for IT service management (BS 15000) was developed by the British Standards Institution (BSI) to enable organizations to enhance the quality of IT services delivered to their customers. The standard specification (BS 15000-1:2002 or Part 1) specifies formal requirements relating to IT service management processes that service providers could implement to assist them in meeting customer and business requirements effectively [BSI 2002]. The specification (Part 1) is to be used in conjunction with Part 2 or the Code of Practice on IT Service Management (BS 15000-2:2003), which provides guidance to auditors assessing service management processes. It also offers guidance and recommendations to organizations planning service improvements or those planning to be audited against BS 15000-1.

BS 15000-1 specifies requirements not only for measuring the quality of services delivered and the effectiveness in meeting business needs at a justifiable cost, but also the performance of service management processes themselves in supporting those objectives. From an overall management perspective (under clause 3.1), BS 15000-1 specifies requirements to define and communicate the service management policy, objectives, and plans, and to communicate the need for meeting the objectives through continual improvement. It also defines requirements for ensuring that customer requirements are met and customer satisfaction is improved. Reviews of service management are required at planned intervals to ensure continuing suitability, adequacy, and effectiveness of the processes [ibid.].

BS 15000-1 includes requirements for a Plan-Do-Check-Act (PDCA) cycle used to plan and implement service management processes.

- Plan service management (Plan)
- Implement service management and provide the services (Do)
- Monitoring, measuring, and reviewing (Check)
- Continuous improvement (Act)

The Check step of the PDCA cycle, corresponding to clause 4.3 of the specification, aims to ensure ongoing control, greater efficiency, and identification of opportunities for continuous improvement. Organizations undergoing an audit for certification are expected to show evidence of control, reporting, and auditing with respect to planning and implementing service management, through monitoring, measurement, and review of not just the services delivered, but also the service management processes. While the focus of Section 4.3 is on achievement of the service management objectives and plan, Section 6.2 focuses on producing agreed-upon, timely, reliable, and accurate reports, on the achievement of service levels, for informed decision making and effective communication on the following [ibid.]:

- performance against service level targets
- non-compliance and issues (e.g., against the SLA, security breach)
- workload characteristics (e.g., volume, resource utilization)
- performance reporting following major events (e.g., major incidents and changes)
- trend information
- satisfaction analysis

While the BS 15000-1 specifies requirements for IT service management at a level suitable for setting objectives, policies, and procedures, detailed process definitions and implementation guidelines are provided by ITIL. The core service management capabilities in ITIL are embodied in the Service Delivery and Service Support processes. For each of the processes ITIL defines the following:

- purpose and objectives to be achieved by the process

- the benefits to be realized from effective implementation of the process
- success factors
- problems and challenges that constitute the risk of failure

Within the context of the purposes, objectives, benefits, and risks for each process, ITIL provides a set of commonly-used performance indicators (KPI) and metrics to achieve visibility and control over the processes. These are to be used by organizations as guidelines for developing measures and indicators suitable to their context.

G.4. Measurement in Software CMM v1.1

The Software CMM provides software organizations with guidance on how to gain control of their processes for developing and maintaining software and how to evolve toward a culture of software engineering and management excellence [Paulk 1995]. The Software CMM was designed to guide software organizations in selecting process improvement strategies by determining current process maturity and identifying the most critical issues for software quality and process improvement. By focusing on a limited set of activities and working aggressively to achieve them, an organization can steadily improve its organization-wide software process to enable continuous and lasting gains in software process capability.

Software process maturity is the extent to which a specific process is explicitly defined, managed, measured, controlled, and effective. Maturity implies a potential for growth in capability and indicates both the richness of an organization's software process and the consistency with which it is applied in projects throughout the organization.

The Software CMM is structured according to several components. Maturity levels (targeted toward building organizational capability), key process areas, and goals are required components. Key practice components are expected, but not required, and subpractices and supplemental information are informative components.

A maturity level is a well-defined evolutionary plateau toward achieving a mature software process. Each maturity level comprises a set of process goals that, when satisfied, stabilize an important component of the software process. Achieving each level of the maturity framework establishes a higher level of process capability for the organization. Organizing the CMM into the five levels shown prioritizes improvement actions for increasing software process maturity.

The five levels can be briefly described as follows:

Table 11
Software CMM Maturity Levels

1) Initial	The software process is characterized as ad hoc, and occasionally even chaotic. Few processes are defined, and success depends on individual effort and heroics.
2) Repeatable	Basic project management processes are established to track cost, schedule, and functionality. The necessary process discipline is in place to repeat earlier successes on projects with similar applications.
3) Defined	The software process for both management and engineering activities is documented, standardized, and integrated into a standard software process for the organization. All projects use an approved, tailored version of the organization's standard software process for developing and maintaining software.
4) Managed	Detailed measures of the software process and product quality are collected. Both the software process and products are quantitatively understood and controlled.
5) Optimizing	Continuous process improvement is enabled by quantitative feedback from the process and from piloting innovative ideas and technologies.

Except for Level 1, each maturity level is decomposed into several key process areas that indicate where an organization should focus to improve its software process. Key process areas identify the issues that must be addressed to achieve a maturity level, as summarized in Table 12.

Table 12
Key Process Areas in the Software CMM

Maturity Level	Focus	Key Process Areas
5 Optimizing	Continual process improvement	Defect Prevention Technology Change Management Process Change Management
4 Managed	Product and process quality	Quantitative Process Management Software Quality Management
3 Defined	Engineering processes and organizational support	Organization Process Focus Organization Process Definition Training Program Integrated Software Management Software Product Engineering Intergroup Coordination Peer Reviews
2 Repeatable	Project management processes	Requirements Management Software Project Planning Software Project Tracking & Oversight Software Subcontract Management Software Quality Assurance Software Configuration Management
1 Initial	Competent people (and heroics)	

Models such as the Software CMM and the eSCM-SP measure organizations or processes against some form of ordinal scale (e.g., maturity levels or capability levels). An alternative improvement strategy is to measure the processes and systems to identify which need to be improved, and use measurement trends to confirm and quantify that improvements are occurring. Measurement-based improvement can be considered part of model-based improvement since measurement-oriented components of the framework usually specify establishing objectives, planning how to achieve those objectives, measuring effectiveness

and efficiency of the processes and systems used, and taking corrective and/or preventive action as appropriate.

Some do not appreciate the degree to which measurement-based improvement is integral to the Software CMM because the model emphasizes implementing good management and engineering practices. Measurement is integral to every key process area in the Software CMM, however. A Measurement and Analysis common feature is part of the Software CMM structure. It describes the need to measure the process and analyze the measurements and typically includes examples of the measurements that could be taken to determine the status and effectiveness of the activities performed in the key process area. Some Measurement and Analysis practices go beyond status:

- the quality of the training program (Training Program)
- the effectiveness of management (Integrated Software Management)
- the functionality and quality of software products (Software Product Engineering)

At Level 2, effective management processes are institutionalized for software projects. They allow organizations to repeat successful practices developed on earlier projects. An effective process can be characterized as practiced, documented, enforced, trained, measured, and able to improve.

At Level 3, common processes, training, and measures are deployed across the organization. A process database containing performance data is established and maintained.

At Level 4, the organization sets quantitative quality goals for software products and processes. Productivity and quality are measured on important software process activities for all projects, as part of an organizational measurement program. Having a defined measurement process in place increases the chances of success by providing a common basis for interpreting measurement. In these cases, data collected from individual projects have significant meaning across projects, so they can significantly increase organizational understanding of the software process. Without a defined measurement process, it is difficult to identify meaningful measurements because of the variation in the processes being measured.

At Level 5, the entire organization is focused on continuous process improvement. The organization has the means to identify weaknesses and strengthen the process proactively, with the goal of preventing the occurrence of defects. Data on the effectiveness of the software process are used to perform cost benefit analyses of new technologies and proposed changes to the organization's software process. Having a defined measurement process in place increases the chances of success by providing a common basis for interpreting measurement. Implementing a managed measurement process provides an understanding of the impact of process changes, thereby increasing the organization's chances of success. By controlling the process within statistically narrow boundaries (small variations in process measures), noise in the data is reduced, making it easier to determine

objectively whether a specific process improvement has an effect. Rational, informed decisions can then be made, based on quantitative foundations.

G.5. Measurement in CMMI v1.1

Capability Maturity Model Integration (CMMI) addresses four disciplines: systems engineering, software engineering, integrated product and process development (IPPD), and supplier sourcing [Chrissis 2003]. Systems engineering covers the development of total systems, which may or may not include software. Systems engineers focus on transforming customer needs, expectations, and constraints into product solutions and supporting these product solutions throughout the life of the product. Software engineering covers the development of software systems. Software engineers focus on applying systematic, disciplined, and quantifiable approaches to the development, operation, and maintenance of software. IPPD is a systematic approach that achieves a timely collaboration of relevant stakeholders throughout the life of the product to better satisfy customer needs, expectations, and requirements. The processes to support an IPPD approach are integrated with the other processes in the organization. As work efforts become more complex, projects may use suppliers to perform functions or add modifications to products that are specifically needed by the project. When those activities are critical, supplier sourcing addresses improved source selection and contract monitoring.

CMMI supports two different architectural representations: staged and continuous. The staged representation uses maturity levels to measure organizational capability, while the continuous representation uses capability levels to measure process capability.

CMMI models are designed to describe discrete levels of process improvement. In the staged representation, five maturity levels provide a recommended order for approaching process improvement. The maturity level of an organization provides a way to predict its future performance within a given discipline or set of disciplines. Experience has shown that organizations do their best when they focus their process-improvement efforts on a manageable number of process areas that require increasingly sophisticated effort as the organization improves. A maturity level is a defined evolutionary plateau of process improvement: Initial, Managed, Defined, Quantitatively Managed, and Optimizing.

Specific goals apply to a process area and address the unique characteristics that describe what must be implemented to satisfy the process area. A practice is an important activity that helps to achieve an associated goal within a process area. Specific practices are expected, but not required, model components.

Generic goals are called "generic" because the same goal statement appears in multiple process areas. In the staged representation, each process area has only one generic goal. In the continuous representation, a process area can be rated at six different capability levels.

Capability levels, which belong to the continuous representation, apply to an organization's process-improvement achievement for each process area. Each of the six capability levels, numbered 0 to 5, corresponds to a generic goal and a set of generic and specific practices.

Process areas are grouped into four categories: Process Management, Project Management, Engineering, and Support. Process Management process areas contain the cross-project activities related to defining, planning, resourcing, deploying, implementing, monitoring, controlling, appraising, measuring, and improving processes. Project Management process areas cover the project management activities related to planning, monitoring, and controlling the project. Engineering process areas cover the development and maintenance activities that are shared across engineering disciplines (e.g., systems engineering and software engineering). Support process areas cover the activities that support product development and maintenance. The Support process areas address processes that are used in the context of performing other processes. The process areas for each category are listed in Table 13.

Table 13
Process Areas in CMMI: SE/SW/IPPD/SS by Process Category

Process Management	Project Management	Engineering	Support
Organizational Process Focus	Project Planning	Requirements Development	Configuration Management
Organizational Process Definition	Project Monitoring and Control	Requirements Management	Process and Product Quality Assurance
Organizational Training	Supplier Agreement Management	Technical Solution	Measurement and Analysis
Organizational Process Performance	Integrated Project Management for IPPD (or Integrated Project Management)	Product Integration	Organizational Environment for Integration
Organizational Innovation and Deployment	Risk Management	Verification	Decision Analysis and Resolution
	Integrated Teaming	Validation	Causal Analysis and Resolution
	Integrated Supplier Management		
	Quantitative Project Management		

At maturity level 2, an organization has achieved all the specific and generic goals of the maturity level 2 process areas. In other words, the organization has ensured that requirements are managed and that processes are planned, performed, measured, and controlled for all projects.

At maturity level 3, processes are managed more proactively using an understanding of the interrelationships of the process activities and detailed measures of the process, its work products, and its services. The organization's measurement repository is used to collect and make available measurement data on processes and work products, particularly as they relate to the organization's set of standard processes. This repository contains or references actual measurement data, along with related information needed to understand and analyze it.

At maturity level 4, quantitative objectives for quality and process performance are established and used as criteria in managing processes. Quantitative objectives are based on the needs of the customer, end users, organization, and process implementers. Quality and process performance are understood in statistical terms and are managed throughout the life of the processes. Quality and process performance measures are incorporated into

the organization's measurement repository to support fact-based decision making in the future.

Maturity level 5 focuses on continually improving process performance through incremental and innovative technological improvements. Quantitative process-improvement objectives for the organization are established, continually revised to reflect changing business objectives, and used as criteria in managing process improvement. The effects of the process improvements are measured and evaluated against the objectives. Both the defined processes and the organization's set of standard processes are targets of measurable improvement activities.

G.6. Measurement in COPC-2000

The COPC-2000 Base Standard [COPC] was developed by the Customer Operation Performance Center Inc. based on the framework of the Malcolm Baldrige National Quality Award [Baldrige]. It provides a set of management practices for "Performance Management" in customer-centric service operations (e.g., customer contact centers, transaction processing centers, fulfillment centers). Measurements and results are key considerations for performance evaluation using this standard.

Category 4.0 (Performance) in COPC-2000 defines the measures and performance expectation from a Customer Service Provider (CSP). Each practice under Category 4.0 defines the following:

- what to measure (with suggested attributes)
- how to set targets
- minimum periodicity of measurements
- recommended sample sizes
- waiver criteria (if any)

The standard prescribes both the Key Customer Related Processes (KCRP) and Key Support Processes (KSP) required for different types of customer service providers (e.g., fulfillment centers and inbound customer contact centers). Additionally, a recommended measure for each process is operationally defined in the appendices (called "exhibits"). Typical measures are listed here:

- On time: for example, percentage of calls answered within a targeted time period
- Backlog: for example, percentage of transactions not processed on time
- Accuracy: for example, defect rate
- Efficiency: for example, average handle time
- Volume: for example, number of calls received per period

A service provider can add more processes and measures to meet the requirements of the customer or business.

The targeted performance levels for each measure are established between the service provider and the customer and documented in a service-level agreement. In order to be certified under the COPC standard, an organization must demonstrate that it is meeting or exceeding the targeted performance level for each of the agreed-to measures and demonstrate sustained improvements over time. There are two criteria for certification or re-certification:

- A minimum of six consecutive months (three months for newly implemented measures) of data are required for first-time certification. A minimum of 12 consecutive months of data are required for re-certification.
- The organization must show sustained improvement. To demonstrate this, the organization must show a minimum of three consecutive data points above its previous performance level for each measure.

The COPC Gold Standard defines more stringent performance requirements for the service provider than the Base Standard. For example, the service provider should meet or exceed targeted performance levels for 65% of its required Category 4.0 performance measures (increased by 15% from the Base Standard) and meet or exceed targeted performance levels or exhibit sustained improvement in 75% of its required Category 4.0 performance measures. This standard also requires measurement of the "cost of poor quality."

Appendix H: Terms and Definitions

Many measurement terms are used inconsistently in the software and IT worlds. For example, "metric" and "measure" are frequently used as synonyms or defined differently. The term "metric" is controversial in software measurement. In this report, the terminology used comes primarily from ISO 15939 (Software Measurement Process) and the eSourcing Capability Model for Service Providers v2

Attribute	A property or characteristic of an entity that can be distinguished quantitatively or qualitatively by human or automated means.
Base measure	A measure defined in terms of an attribute, and the method for quantifying it. (Based on International Vocabulary of Basic and General Terms in Metrology 1993.)
	Note: A base measure is functionally independent of other measures.
Benchmark	A reference point or standard by which something can be judged. (Adapted from [Camp 1989].)
Benchmarking	1) The search for industry best practices that lead to superior performance. 2) The continuous process of measuring products, services, and practices against competitors or industry leaders.
Derived measure	A measure that is defined as a function of two or more values of base measures.
End-user	The ultimate consumer of services provided by the service provider or client. For example, in the case of a software company sourcing its customer service call center, end-users are the customers who call into the call center run by the service provider. End-users may be part of the client organization, or may be customers of the client.
Engagement	The relationship between the service provider and a current or prospective client that spans the entire Sourcing Life-cycle.
Engagement objectives	A set of goals that are used to set direction in the sourcing engagement. Engagement objectives should be measurable so that progress against them can be tracked. They are typically based on client relationship factors (for example, increase client satisfaction by x%) or on operational factors (for example, improve productivity by y%).
Entity	An object that is to be characterized by measuring its attributes. An entity can be a process, product, project, or resource.
Indicator	A measure that provides an estimate or evaluation of specified attributes derived from a model, with respect to defined information needs.
Measure (noun)	A variable to which a value is assigned as the result of measurement.
	Note: The term "measures" is used to refer collectively to base measures, derived measures, and indicators.
Measure (verb)	To make a measurement [ISO 1996].
Measurement	A set of operations having the object of determining a value of a measure.
Nonconformance	A failure to satisfy a requirement, which may be specified in a policy, procedure, standard, statute, regulation, service level agreement, or contract.
Organization	As used in the eSCM-SP Practices, an organization is an entity that provides sourcing services to one or more clients. Depending on its size or complexity, a single company may have one or more service provision organizations.

Organizational objectives	A formal set of objective or quantified business goals that are used to set long-term direction. Examples of organizational objectives include increasing client satisfaction by 5% based on feedback forms, maintaining client satisfaction, growing market share by 8%, and improving performance by 12%.
Peer group	A set of similar entities that can be considered a homogenous sample for a population being statistically analyzed
Peer group criteria	The characteristics that determine similarity of entities in a peer group. Note: The peer group criteria for a service provider includes items such as market sector, service, size of service provider (or relevant organizational unit), complexity of service, and geographical location.
Process	1) A set of actions that is performed to achieve a given purpose, along with the assets that support that performance, such as tools and other resources. 2) A set of interrelated activities that transform inputs into outputs [ISO 1998].
Process capability	The range of expected results that can be achieved by following a process. The process capability of an organization provides one means of predicting the most likely outcomes to be expected from the next project the organization undertakes [Paulk 1995].
Resources	1) Resources include all of the following: people, skills, experience, knowledge assets, intellectual property, processes and guidelines, repository, solutions, documents, infrastructure, computers, storage, networks, data, applications, facilities, financial [IEEE-STD-610 1990]. 2) The assets available for providing a product or service, including people, knowledge assets, infrastructure, and finances. Note 1: People resources may be characterized by skills and experience. Note 2: Knowledge assets include intellectual property, processes, solutions, and documents. Note 3: Infrastructure includes facilities, networks, computers, applications, and other aspects of the work environment.
User	As used in this report, users are the end-users, consumers of the services provided by the service providers.