AN ASSESSMENT MODEL FOR THE APPLICABILITY OF STATISTICAL PROCESS CONTROL FOR SOFTWARE PROCESSES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYÇA TARHAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

OCTOBER 2006

Approval of the Graduate School of Informatics

____

Assoc.Prof.Dr. Nazife BAYKAL

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

____

Assoc.Prof.Dr. Yasemin YARDIMCI

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

____

Assoc.Prof.Dr. Onur DEMİRÖRS

Supervisor

Examining Committee Members

Prof.Dr. Semih BİLGEN (METU, EEE) __________________________

Assoc.Prof.Dr. Onur DEMİRÖRS (METU, IS) __________________________

Dr. Ali ARİFOĞLU (METU, IS) __________________________

Assoc.Prof.Dr. Ali DOĞRU (METU, CENG) __________________________

Assist.Prof.Dr. Y. Murat ERTEN (TOBB ETU, CENG) __________________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Ayça Tarhan

Signature: ________________

ABSTRACT

AN ASSESSMENT MODEL FOR THE APPLICABILITY OF
STATISTICAL PROCESS CONTROL FOR SOFTWARE PROCESSES

Tarhan, Ayça
Ph.D., Department of Information Systems
Supervisor: Assoc. Prof. Dr. Onur Demirörs

October 2006, 264 pages

The application of statistical process control (SPC) techniques for software is rare due to such requirements as high maturity, rational sampling, and effective metric selection. Companies that invest time and money on a process improvement model can also take the advantage of following a well-founded framework to establish the infrastructure required for SPC implementation. For other companies, however, the path to SPC implementation is not that clear. Existing studies frequently focus on potential benefits of SPC results rather than providing guidelines based on practical evidence.

In this study, we developed an assessment model to test the applicability of SPC for software processes, and performed three case studies in a multiple-case-study context to answer two basic questions: 1) Can we identify guidelines to direct SPC implementation? 2) Can emergent organizations apply SPC techniques following these guidelines and benefit from the results? We worked on task management, review, test development processes and related metrics of different organizations. As control chart is one of the most sophisticated data analysis tools within SPC, we demonstrated practical evidence on the utilization of SPC via control charts. Multiple case study results showed us that with established guidelines for rational sampling and metric utilization, emergent organizations can apply SPC techniques and attain the ability to understand its processes based on quantitative data.

Key Words: Statistical process control, rational sampling, measurement, control chart.

ÖZ

İSTATİSTİKSEL SÜREÇ KONTROLÜNÜN YAZILIM SÜREÇLERİNE UYGULANABİLİRLİĞİ İÇİN BİR DEĞERLENDİRME MODELİ

Tarhan, Ayça
Doktora, Bilişim Sistemleri Bölümü
Tez Yöneticisi: Doç. Dr. Onur Demirörs

Ekim 2006, 264 sayfa

Yazılım için İstatistiksel Süreç Kontrolü (İSK) uygulamaları; olgunluk seviyesinin yüksekliği, rasyonel örnekleme ve metriklerin etkin seçimi gibi gereksinimler sebebiyle enderdir. Bir süreç iyileştirme modeline zaman ve kaynak ayıran firmalar, modelin İSK uygulamaları için gerekli altyapının kurulmasını destekleyen iyi tanımlı çatısını izlemekten yararlanabilirler. Ancak diğer firmalar için İSK uygulamalarını başarmaya giden yol, o kadar net değildir. Mevcut çalışmalar, pratik kanıtlara dayanarak kılavuzluk etmek yerine, genellikle İSK sonuçlarının potansiyel getirileri üzerine odaklanmıştır.

Biz bu çalışmada, İSK'nın yazılım süreçlerine uygulanabilirliğini test etmek için bir değerlendirme modeli geliştirdik ve şu iki temel soruyu yanıtlamak üzere üç örnek çalışma yaptık: 1) İSK uygulamalarına kılavuzluk edecek bir yöntem tanımlayabilir miyiz? 2) Gelişmekte olan kurumlar bu yöntemi uygulayabilir ve sonuçlarından fayda sağlayabilirler mi? Çalışmalarda farklı kurumların görev yönetimi, gözden geçirme ve test geliştirme süreçleri ve ilişkili metrikleri üzerinde çalıştık. Kontrol grafikleri İSK seti içindeki gelişmiş veri analiz araçlarından biri olduğundan, İSK'nın kullanılabilirliğine dair pratik kanıtı kontrol grafikleri ile gösterdik. Çoklu çalışma sonuçları bize; gelişmekte olan kurumların, rasyonel örnekleme ve metrik kullanılabilirliğine dair tanımlı yöntemleri izleyerek İSK tekniklerini kullanabildiğini ve süreçlerini nicel veriye dayalı olarak anlama yetkinliğini kazanabildiğini kanıtladı.

Anahtar Kelimeler: İstatistiksel süreç kontrolü, rasyonel örnekleme, ölçme, kontrol grafiği

To the masters in my life,

who taught me many

ACKNOWLEDGEMENTS

I send my deepest regards and gratitude to Assoc. Prof. Onur Demirörs for his expertise, insight, and guidance, and for his patience and support during my doctorate study.

I express my sincere gratitude to Prof. Semih Bilgen and Assoc. Prof. Ali Doğru for their insight and feedback throughout the progress of my study.

I am grateful to E.Y., M.S.T., and M.S. for their help and contribution during case study implementations. I am also grateful to my colleagues at my workplace for their fun and support.

I send my warm regards and thanks to my parents for always trusting and supporting me in whatever I do.

I have the deepest emotions for Özgür, for his enduring patience and support, and mostly for sharing the life with me.

And finally, I appreciate myself for understanding the pleasures and difficulties of this long-lasting period and for not losing my faith that these studies would come to an end.

TABLE OF CONTENTS

ABSTRACT iv

ÖZ v

DEDICATION vi

ACKNOWLEDGEMENTS vii

TABLE OF CONTENTS viii

LIST OF TABLES x

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xxv

CHAPTER

1. INTRODUCTION 1

1.1. Utilization of SPC for Software Processes 2

1.2. An Assessment Model for Statistical Process Control (SPC-AM) . 3

1.3. Roadmap 7

2. BACKGROUND 9

2.1. SPC Concepts and Tools 9

2.1.1 Tools Used for SPC 11

2.1.2 Shewhart's Control Charts 12

2.2. Literature on SPC for Software 18

2.2.1 Implementations of SPC for Software 18

2.2.2 Guidance on SPC for Software 26

2.3. Literature on Measurement Approaches and Models 36

2.4. Literature on Measurement Practices 45

3. AN ASSESSMENT MODEL FOR STATISTICAL PROCESS CONTROL ..
.... 57

3.1. Model Components 57

3.2. Assessment Process 61

3.3. Assessment Assets 64

3.4. Relation of the Model with the CMMI's Measurement Practices 71

4. APPLICATION OF THE MODEL 73

4.1. Fundamentals of Design 73

4.2. Multiple-Case Study Design 76

4.3. Context-1 (Case Study-A) 78

4.4. Context-2 (Case Study-B) 87

4.5. Context-3 (Case Studies C, D, and E) 108

5. CONCLUSION 123

REFERENCES … 131

APPENDICES

A. ASSESSMENT ASSETS 136

B. CASE STUDY PLANS 143

C. DETAILS OF CASE STUDY-A 150

D. DETAILS OF CASE STUDY-B 168

E. DETAILS OF CASE STUDIES C, D, E 204

CURRICULUM VITAE … 263

LIST OF TABLES

Table 2.1 Crucial Points on SPC for Software 33

Table 2.2 Capability Levels and Process Attributes in ISO/IEC TR 15504......... 45

Table 2.3 CMMI Measurement and Analysis Process Area – Goals and Practices 48

Table 2.4 CMMI Quantitative Project Management Process Area – Goals and Practices 49

Table 3.1 Metric Usability Attributes used for Evaluating Metric Utilization 60

Table 3.2 Model’s Traceability to CMMI’s Measurement Practices 72

Table 4.1 Initial Results from Charted Data for Task Management Process Derived Metrics 84

Table 4.2 Assignable Causes for Out-of-Control Points of Task Mgt. Process.... 85

Table 4.3 Usability Evaluation Results for Review Process Metrics 95

Table 4.4 Sub-Clusters of M with respect to Input Product Types 96

Table 4.5 Initial Results from Charted Data for Review Process Derived Metrics 97

Table 4.6 Assignable Causes for Out-of-Control Points of Review Process...... 102

Table 4.7 Final Results from Re-charted Data for Review Process Derived Metrics 103

Table 4.8 Processes and Process Clusters Identified for Cases C, D, E (Context-3) 111

Table 4.9 Processes and Data Sets (Original and Derived) in Context-3........... 112

Table 4.10 Metric Usability Results in Context-3 115

Table 4.11 Derived Metrics Utilized to Understand Process Performances in Context-3 115

Table 4.12 Initial Results from Charted Data in Context-3 117

Table 4.13 Assignable Causes for Out-of-Control Points in Context-3 119

Table 4.14 Final Results from Charted Data in Context-3 120

Table 4.15 Mean Values of Process Metrics for Each Cluster in Context-3 121

Table B.1 Work Breakdown Structure for Case Studies 149

LIST OF FIGURES

Figure 1.1 Factors That Contribute To Starting SPC Implementation 5

Figure 2.1 Example Control Chart 14

Figure 2.2 The Goal-Question-Metric Hierarchy 38

Figure 2.3 The Model Measure Manage Paradigm 41

Figure 2.4 The Measurement Process Model of PSM 44

Figure 2.5 ISO/IEC 15939 Software Measurement Process 52

Figure 3.1 Process Attributes used for Rational Sampling 58

Figure 3.2 The Assessment Process 61

Figure 3.3 Process Execution Record 64

Figure 3.4 Process Similarity Matrix 65

Figure 3.5 Process Clusters and Cluster Distances 66

Figure 3.6 Metric Usability Questionnaire and Rating for Base Metrics 68

Figure 3.7 Metric Usability Questionnaire and Rating for Derived Metrics 69

Figure 3.8 Process Execution Questionnaire 70

Figure 4.1 The Structure of Embedded, Multiple-Case Design 75

Figure 4.2 Task Management Process States 78

Figure 4.3 Process Execution Record Used to Sample Task Management Process Executions 79

Figure 4.4 Process Similarity Matrix for (the first 20) Task Management Process Executions 80

Figure 4.5 Initial Process Clusters and Sub-clusters for Task Management Process 81

Figure 4.6 Task Management Base and Derived Metrics.................................. 81

Figure 4.7 Individuals Chart for Combined Data (Clusters C and D) of Effort Variance 82

Figure 4.8 Individuals Charts for Effort Variance 83

Figure 4.9 Individuals Charts for Effort Variance 83

Figure 4.10 Process Execution Questionnaire for Task Mgt. Process Execution # 1 84

Figure 4.11 Review Process as Defined in the Company-Specific Procedure 88

Figure 4.12 Process Execution Record Used to Sample Review Process Executions 89

Figure 4.13 Process Similarity Matrix for (the first 20) Review Process Executions 89

Figure 4.14 Initial Process Clusters and Cluster Distances for Review Process .. 90

Figure 4.15 Review Base and Derived Metrics 92

Figure 4.16 Metric Usability Questionnaire for “Review Effort” Base Metric 93

Figure 4.17 Metric Usability Questionnaire for “Nonconformance Detection Efficiency” Derived Metric 94

Figure 4.18 Individuals Charts for Nonconformance Detection Efficiency......... 96

Figure 4.19 Individuals Charts for Base Metrics of Nonconformance Detection Efficiency for Process Cluster M-1 98

Figure 4.20 Individuals Charts for Base Metrics of Nonconformance Resolution Efficiency for Process Cluster M-2 99

Figure 4.21 Individuals Chart for “Number of Accepted Nonconformances” Base Metric of Open Period with respect to Nonconformances for Process Cluster M-3.. 99

Figure 4.22 Process Execution Questionnaire for Review Process Execution # 98 100

Figure 4.23 Re-charting of Nonconformance Detection Efficiency for Cluster M-1 103

Figure 4.24 Review Process Flow after Case Study Implementation 107

Figure 4.25 Flow of Test Design Process 109

Figure 4.26 Process Similarity Matrix for Test Design 110

Figure 4.27 Initial Process Clusters for Test Design Process 111

Figure 4.28 Base and Derived Metrics Identified in Context-3 113

Figure 4.29 Metric Usability Questionnaire for "Actual Test Design Effort" Base Metric of "Test Design" Process 114

Figure 4.30 Individuals Chart for Combined Data of Test Design Productivity 116

Figure 4.31 Individuals Charts for Test Design Productivity 116

Figure 4.32 Process Execution Questionnaire for Test Procedure Development Process Execution # 1 118

Figure A.1 Process Execution Record 137

Figure A.2 Process Execution Questionnaire 138

Figure A.3 Process Similarity Matrix 139

Figure A.4 Process Attributes Description 140

Figure A.5 Metric Usability Questionnaire for Base Metrics 141

Figure A.6 Metric Usability Questionnaire for Derived Metrics 142

Figure B.1 Work Plan for Case Study-A 144

Figure B.2 Work Plan for Case Study-B 146

Figure B.3 Work Plan for Case Studies C, D, and E 148

Figure C.1 Process Execution Record for Task Management PE # 5 150

Figure C.2 Process Execution Record for Task Management PE # 10 151

Figure C.3 Process Execution Record for Task Management PE # 25 151

Figure C.4 Process Execution Record for Task Management PE # 57 152

Figure C.5 Process Similarity Matrix for Task Management Process 153

Figure C.6 Metric Usability Questionnaire for “Estimated Start Date” Base Metric of Task Management Process 154

Figure C.7 Metric Usability Questionnaire for “Estimated Finish Date” Base Metric of Task Management Process.. 155

Figure C.8 Metric Usability Questionnaire for “Actual Start Date” Base Metric of Task Management Process 156

Figure C.9 Metric Usability Questionnaire for “Actual Finish Date” Base Metric of Task Management Process 157

Figure C.10 Process Execution Questionnaire for Task Management PE # 1.... 158

Figure C.11 Process Execution Questionnaire for Task Management PE # 26.. 158

Figure C.12 Process Execution Questionnaire for Task Management PE # 34.. 159

Figure C.13 Process Execution Questionnaire for Task Management PE # 49.. 159

Figure C.14 Process Execution Questionnaire for Task Management PE # 79.. 159

Figure C.15 Process Attributes Description for Process Cluster C2................. 160

Figure C.16 Process Attributes Description for Process Cluster C-except C2 ... 161

Figure C.17 Process Attributes Description for Process Cluster D................... 162

Figure C.18 Control Chart for Estimation Capability of Overall Task Management Process 163

Figure C.19 Control Chart for Estimation Capability of Task Management Process Cluster C.. 163

Figure C.20 Control Chart for Estimation Capability of Task Management Process Cluster C2.. 164

Figure C.21 Control Chart for Estimation Capability of Task Management Process Cluster C-except C2.. 164

Figure C.22 Control Chart for Estimation Capability of Task Management Process Cluster D 165

Figure C.23 Control Chart for Effort Variance of Overall Task Management Process 165

Figure C.24 Control Chart for Effort Variance of Task Management Process Cluster C... 166

Figure C.25 Control Chart for Effort Variance of Task Management Process Cluster C2... 166

Figure C.26 Control Chart for Effort Variance of Task Management Process Cluster C-except C2... 167

Figure C.27 Control Chart for Effort Variance of Task Management Process Cluster D 167

Figure D.1 Process Execution Record for Review PE # 10.............................. 168

Figure D.2 Process Execution Record for Review PE # 30.............................. 169

Figure D.3 Process Execution Record for Review PE # 53.............................. 169

Figure D.4 Process Execution Record for Review PE # 177............................ 170

Figure D.5 Process Execution Record for Review PE # 189............................ 170

Figure D.6 Process Similarity Matrix for Review Process............................... 171

Figure D.7 Metric Usability Questionnaire for "Opening Date" Base Metric of Review Process... 173

Figure D.8 Metric Usability Questionnaire for "Closure Date" Base Metric of Review Process... 174

Figure D.9 Metric Usability Questionnaire for "Number of Detected Nonconformances" Base Metric of Review Process 175

Figure D.10 Metric Usability Questionnaire for "Number of Accepted Nonconformances" Base Metric of Review Process 176

Figure D.11 Metric Usability Questionnaire for "Review Effort" Base Metric of Review Process... 177

Figure D.12 Metric Usability Questionnaire for "Nonconformance Resolution Effort" Base Metric of Review Process ... 178

Figure D.13 Metric Usability Questionnaire for "Nonconformance Detection Efficiency" Derived Metric of Review Process ... 179

Figure D.14 Metric Usability Questionnaire for "Nonconformance Resolution Efficiency" Derived Metric of Review Process ... 180

Figure D.15 Metric Usability Questionnaire for "Open Period" Derived Metric of Review Process ... 181

Figure D.16 Metric Usability Questionnaire for "Open Period with respect to Nonconformances" Derived Metric of Review Process ... 182

Figure D.17 Process Execution Questionnaire for Review PE # 51 ... 183

Figure D.18 Process Execution Questionnaire for Review PEs # 59, 62, 63, 64 183

Figure D.19 Process Execution Questionnaire for Review PE # 70 ... 184

Figure D.20 Process Execution Questionnaire for Review PE # 98 ... 184

Figure D.21 Process Execution Questionnaire for Review PE # 156 ... 184

Figure D.22 Process Attributes Description for Review Process Cluster M ... 185

Figure D.23 Process Attributes Description for Review Process Cluster N ... 186

Figure D.24 Control Chart for Nonconformance Detection Efficiency of Overall Review Process ... 187

Figure D.25 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M ... 187

Figure D.26 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-1 ... 188

Figure D.27 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-2 ... 188

Figure D.28 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-3 ... 189

Figure D.29 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-4 189

Figure D.30 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster N 190

Figure D.31 Control Chart for Nonconformance Resolution Efficiency of Overall Review Process 191

Figure D.32 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M 191

Figure D.33 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-1 192

Figure D.34 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-2 192

Figure D.35 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-3 193

Figure D.36 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-4 193

Figure D.37 Control Chart for Open Period with respect to Nonconformances of Overall Review Process 194

Figure D.38 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M 194

Figure D.39 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-1 195

Figure D.40 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-2 195

Figure D.41 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-3 196

Figure D.42 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-4 196

Figure D.43 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster N 197

Figure D.44 Control Chart for Open Period of Overall Review Process 198

Figure D.45 Control Chart for Open Period of Review Process Cluster M 198

Figure D.46 Control Chart for Open Period of Review Process Cluster M-1 199

Figure D.47 Control Chart for Open Period of Review Process Cluster M-2 199

Figure D.48 Control Chart for Open Period of Review Process Cluster M-3 200

Figure D.49 Control Chart for Open Period of Review Process Cluster M-4 200

Figure D.50 Control Chart for Open Period of Review Process Cluster N 201

Figure D.51 Control Chart for Defect Density of Overall Review Process 202

Figure D.52 Control Chart for Defect Density of Review Process Cluster M ... 202

Figure D.53 Control Chart for Defect Density of Review Process Cluster N 203

Figure E.1 Test Design Process Flow 205

Figure E.2 Test Procedure Development Process Flow 206

Figure E.3 Test Development Peer Review Process Flow 207

Figure E.4 Process Similarity Matrix for Test Design Process 209

Figure E.5 Process Similarity Matrix for Test Procedure Development Process 210

Figure E.6 Process Similarity Matrix for Test Development Peer Review Process 211

Figure E.7 Metric Usability Questionnaire for “Test Design Internal Review Effort” Base Metric of Test Design Process 213

Figure E.8 Metric Usability Questionnaire for “Number of Test Cases” Base Metric of Test Design Process 214

Figure E.9 Metric Usability Questionnaire for “Actual Test Design Effort” Base Metric of Test Design Process 215

Figure E.10 Metric Usability Questionnaire for "Test Design Productivity" Derived Metric of Test Design Process 216

Figure E.11 Metric Usability Questionnaire for "Percent of Test Design Internal Review Effort" Derived Metric of Test Design Process 217

Figure E.12 Metric Usability Questionnaire for "Actual Test Procedure Development Effort" Base Metric of Test Procedure Development Process 218

Figure E.13 Metric Usability Questionnaire for "Test Procedure Development Internal Review Effort" Base Metric of Test Procedure Development Process 219

Figure E.14 Metric Usability Questionnaire for "Percent of Test Procedure Development Internal Review Effort" Derived Metric of Test Procedure Development Process 220

Figure E.15 Metric Usability Questionnaire for "Test Procedure Development Productivity" Derived Metric of Test Procedure Development Process 221

Figure E.16 Metric Usability Questionnaire for "Actual Test Development Effort" Derived Metric of Test Development 222

Figure E.17 Metric Usability Questionnaire for "Percent of Test Development Internal Review Effort" Derived Metric of Test Development.................. 223

Figure E.18 Metric Usability Questionnaire for "Test Development Internal Review Effort" Derived Metric of Test Development............................... 224

Figure E.19 Metric Usability Questionnaire for "Test Development Productivity" Derived Metric of Test Development 225

Figure E.20 Metric Usability Questionnaire for "Number of Action Items" Base Metric of Test Development Peer Review Process 226

Figure E.21 Metric Usability Questionnaire for "Test Development Peer Review Effort" Base Metric of Test Development Peer Review Process 227

Figure E.22 Metric Usability Questionnaire for "Test Development Peer Review Update Effort" Base Metric of Test Development Peer Review Process ... 228

Figure E.23 Metric Usability Questionnaire for “Action Item Density” Derived Metric of Test Development Peer Review Process 229
Figure E.24 Metric Usability Questionnaire for “Action Item Detection Efficiency” Derived Metric of Test Development Peer Review Process ... 230
Figure E.25 Metric Usability Questionnaire for “Action Item Resolution Efficiency” Derived Metric of Test Development Peer Review Process ... 231
Figure E.26 Metric Usability Questionnaire for “Total Review Effort” Derived Metric of Overall Reviews 232
Figure E.27 Metric Usability Questionnaire for “Total Review Effort per Test Case” Derived Metric of Overall Reviews... 233
Figure E.28 Process Execution Questionnaire for Test Design PE # 11............ 234
Figure E.29 Process Execution Questionnaire for Test Proc. Development PE # 1 235
Figure E.30 Process Execution Questionnaire for Test Proc. Development PE # 11 235
Figure E.31 Process Execution Questionnaire for Test Proc. Development PE # 45 235
Figure E.32 Process Execution Questionnaire for Test Development PE # 11 .. 236
Figure E.33 Process Execution Questionnaire for Test Development PE # 45 .. 236
Figure E.34 Process Execution Questionnaire for Test Development Peer Review PE # 44 236
Figure E.35 Process Execution Questionnaire for Overall Reviews PE # 14..... 237
Figure E.36 Process Execution Questionnaire for Overall Reviews PE # 24..... 237
Figure E.37 Process Attributes Description for Test Design Process Cluster 1 . 239
Figure E.38 Process Attributes Description for Test Design Process Cluster 2 . 240
Figure E.39 Process Attributes Description for Test Procedure Development Process 241

Figure E.40 Process Attributes Description for Test Development Peer Review Process 242

Figure E.41 Control Chart for Productivity of Test Design Process (Overall)... 243

Figure E.42 Control Chart for Productivity of Test Design Process Cluster-1... 243

Figure E.43 Control Chart for Productivity of Test Design Process Cluster-2... 244

Figure E.44 Control Chart for Percent of Test Design Internal Review Effort of Test Design Process (Overall) 245

Figure E.45 Control Chart for Percent of Test Design Internal Review Effort of Test Design Process Cluster-1 245

Figure E.46 Control Chart for Percent of Test Design Internal Review Effort of Test Design Process Cluster-2 246

Figure E.47 Control Chart for Productivity of Test Procedure Development Process (Overall) 247

Figure E.48 Control Chart for Productivity of Test Procedure Development Process Cluster-1 247

Figure E.49 Control Chart for Productivity of Test Procedure Development Process Cluster-2 248

Figure E.50 Control Chart for Productivity of Test Procedure Development Process Cluster-3 248

Figure E.51 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process (Overall)....................................... 249

Figure E.52 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process Cluster-1 249

Figure E.53 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process Cluster-2 250

Figure E.54 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process Cluster-3 250

Figure E.55 Control Chart for Productivity of Test Development (Overall)...... 251

Figure E.56 Control Chart for Productivity of Test Development Cluster-1 251

Figure E.57 Control Chart for Productivity of Test Development Cluster-2 252

Figure E.58 Control Chart for Productivity of Test Development Cluster-3 252

Figure E.59 Control Chart for Percent of Test Design Internal Review Effort of Test Development (Overall) 253

Figure E.60 Control Chart for Percent of Test Design Internal Review Effort of Test Development Cluster-1 253

Figure E.61 Control Chart for Percent of Test Design Internal Review Effort of Test Development Cluster-2 254

Figure E.62 Control Chart for Percent of Test Design Internal Review Effort of Test Development Cluster-3 254

Figure E.63 Control Chart for Action Item Density of Test Development Peer Review Process (Overall) 255

Figure E.64 Control Chart for Action Item Density of Test Development Peer Review Process Cluster-1 255

Figure E.65 Control Chart for Action Item Density of Test Development Peer Review Process Cluster-2 256

Figure E.66 Control Chart for Action Item Density of Test Development Peer Review Process Cluster-3 256

Figure E.67 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process (Overall) 257

Figure E.68 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process Cluster-1 257

Figure E.69 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process Cluster-2 258

Figure E.70 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process Cluster-3 258

Figure E.71 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process (Overall) 259

Figure E.72 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process Cluster-1 259

Figure E.73 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process Cluster-2 260

Figure E.74 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process Cluster-3 260

Figure E.75 Control Chart for Overall Review Effort per Test Case of Overall Reviews (Overall) 261

Figure E.76 Control Chart for Overall Review Effort per Test Case of Overall Reviews Cluster-1 261

Figure E.77 Control Chart for Overall Review Effort per Test Case of Overall Reviews Cluster-2 262

Figure E.78 Control Chart for Overall Review Effort per Test Case of Overall Reviews Cluster-3 262

LIST OF ABBREVIATIONS

AB: Ability to Perform

CSCI: Computer Software Configuration Item

CMM: Capability Maturity Model

CMMI: Capability Maturity Model Integrated

CMU: Carnegie Mellon University

CO: Commitment to Perform

DI: Directing Implementation

eEPC: Extended Event-Driven Process Chain

GG: Generic Goal

GP: Generic Practice

GQM: Goal-Question-Metric

GQIM: Goal-Question-Indicator-Measure

IDD: Interface Design Description

IEC: International Electrotechnical Commission

IRS: Interface Requirements Specification

ISO: International Standards Organization

İSK: İstatistiksel Süreç Kontrolü

L3: Maturity Level 3

L5: Maturity Level 5

LCL: Lower Control Limit

LOC: Lines of Code

M3P: Model Manage Measure Paradigm

MUF: Metric Usability Factor

MUQ: Metric Usability Questionnaire

OCP: Out-of-Control Point

PEQ: Process Execution Questionnaire

PSM: Practical Software Measurement

QPM: Quantitative Process Management

SDD: Software Design Description

SEI: Software Engineering Institute

SEPG: Software Engineering Process Group

SG: Specific Goal

SP: Specific Practice

SLOC: Source Lines of Code

SPC: Statistical Process Control

SPC-AM: Assessment Model for Statistical Process Control

SRS: Software Requirements Specification

SW: Software

TL: Team Leader

UCL: Upper Control Limit

VE: Verification

WG: Workgroup

XmR: X (Individual) and Moving Range

CHAPTER 1

INTRODUCTION

Measurement is vital in any engineering discipline, and software measurement is not an exception. Considering the saying "An engineering discipline is as mature as its measurement tools." by Louis Pasteur, we can propose that the state of the practice for software measurement shows that software engineering is somewhere in the middle of its maturing process. Measurement in software industry has been considered a luxury for many years [1]. As the competition has escalated among software development organizations for supplying high quality, timely, and less costly software to their customers; structured measurement programs has started to gain attention in the sector. Many organizations are seeking ways to start a formal measurement program or apply basic measurement practices in the context of process improvement models like CMMI [2] or ISO/IEC 15504 [3], in order to incorporate product and process measures into their planning and decision making processes. By doing so, they intend to gain the control of their processes at all levels.

Attaining and maintaining software process control can be easily proposed but it is proven to be difficult-in practice. The method that has been most widely used in manufacturing domains for this purpose is the Statistical Process Control (SPC). SPC is a powerful collection of problem solving tools that are used for achieving process stability and improving process capability through the reduction of variability [4]. It was first proposed by Shewhart in 1930s [5] and sophisticated by Deming's studies [6][7].

While benefits of SPC are proven for manufacturing companies, they have not yet been in software domain. Software development differs from manufacturing in many ways. First, people are inseparable components of software development. Second, transformation of user requirements into software is dominated by cognitive activities [8]. Third, software development process does not involve repeated delivery of equivalent services or the fabrication of identical products [9]. We need to adapt the concepts of statistical process control for application in software industry, and we need to set the ground for its successful application. Setting the ground includes providing guidelines for efficient metric definition, reliable data collection, and effective SPC analysis, at a minimum.

1.1. Utilization of SPC for Software Processes

Little number of studies has been reported on SPC implementation for software. The earliest study was presented in 1999, explaining the results from a cooperative effort where Software Engineering Institute and the Space Shuttle Onboard Software Project experiment applying SPC analysis to inspection activities [10]. Florac et. al. described the experiences of SPC implementations on the same project [11] and stated a number of notions to consider for fully appreciating how control charts are used to measure and analyze software processes. In year 2000, Weller provided a distinct case in an article by explaining details on SPC implementation to analyze inspection and test data in a software organization [12]. Another study was reported in the same year by Jalote, Dinesh, Raghavan, Bhashyam, and Ramakrishnan [13], which described the approach of quantitative quality management through defect prediction and SPC employed at Infosys. Jacob and Pillai explained details related to SPC implementation via control charts to control variation in the coding and code review processes in an article published in 2003 [14]. Another implementation was reported by Demirors and Sargut in 2003 [15]. They described the difficulties and suggestions in application of SPC to a CMM Level 3 organization using defect density metric. Based on this implementation, the authors later published pitfalls and suggestions of utilizing statistical process control in emergent organizations in another study [16].

Aside from specific reports of SPC implementations for software processes summarized above, the literature holds a number of articles and tutorials that discuss the reasons of difficulties and provide suggestions on the subject. [8][9][17][18][19][20][21]. Lantzy and Card attribute scarcity of SPC implementations for software to the inherent properties of the domain. Lantzy states that transformation of user requirements into software is dominated by cognitive activities and higher-level cognition increases variances in productivity and quality, making application of SPC difficult [8]. Card claims that software development process does not involve repeated delivery of equivalent services or the fabrication of identical products, meaning that process variation is natural [9]. Radice, on the other hand, argues that there are software processes such as configuration management, planning, estimating, tracking, defect prevention, and inspection, for which above difficulties are less prominent [21].

As process improvement models like CMM [22], ISO/IEC 15504 [3] or CMMI [2] have become popular during the last decade, SPC for software has gained attention. These models implicitly direct companies to implement SPC as a crucial step for achieving higher maturity levels [23][24]. Once a company invests on one of these models, it can also take the advantage of following a well-founded framework to establish the infrastructure required for SPC implementation. For other companies, however, the path to SPC implementation is not that clear. Existing implementations focus on the potential benefits of SPC results rather than on providing satisfactory guidelines based on practical evidence. We lack knowledge on the techniques for rational sampling and sub-grouping, applicability of different metrics, the means of reliable data collection mechanisms, and meaningful data analysis, especially for emergent organizations.

1.2. An Assessment Model for Statistical Process Control (SPC-AM)

The need for the knowledge mentioned above encouraged us develop an assessment model to evaluate the suitability of SPC for software process and metrics. Accordingly, we intended to answer two basic questions throughout our research study: 1) Can we identify guidelines to direct SPC implementation? 2)

Can emergent organizations apply SPC techniques following these guidelines and benefit from the results?

To identify the guidelines, we should have first clarified the problems or obstacles to SPC implementation. Therefore we first focused on investigating the challenges that might hinder the implementation for software, which would in turn show us the issues to address by our study.

One of the challenges is related with management. Due to its inherent characteristics as people-dependency, product invisibility and changeability, the software domain has been suffering from the lack of effective control loops based on quantitative data at many levels. Implementing such a control loop at organizational level was reported as being not easy and requiring hard work for many years [25][26][27]. If the business goals are not aligned with the goals of process understanding and improvement (specifically with the targets of SPC here), the motivation for measurement and analysis cannot be initiated and/or maintained because the use of results to be generated cannot be understood by process stakeholders.

The earliest and most investigated approach for goal-based measurement is the Goal-Question-Metric (GQM) paradigm [28] which proposes a top-down style of measurement definition. The approach states that for an organization to measure in a purposeful way, it must specify the goals for itself and its projects; it must trace those goals to the data that are intended to define the goals operationally, and provide a framework for interpreting the data with respect to the stated goals [29]. Other models have been developed based on GQM, including Goal-Question-Indicator-Measure (GQIM) model [30] and Model Measure Manage Paradigm (M^3P) [31]. An issue with top-down style of measurement is that it somehow cannot meet the bottom-up nature of software process improvement. Since software development requires knowledge work and its integration, they are the practitioners at the bottom of the organization that should define, measure, and improve their processes [32][33]. Some pragmatic approaches to software measurement definition and data collection have also been proposed as opposed to the top-down, goal-driven measurement, such as the bottom-up approach that states organizations should measure what is available regardless of goals [34]. Although the idea seems questionable at first for the purpose it serves, it might be a solution for emergent organizations.

The issues regarding the management challenge mentioned above largely depend on the culture and habits of an organization and are not easy to address in the short term. We believe this is one of the contributors to scarce number of SPC implementations for software. Organizational maturity as stated by CMM and CMMI has a supporting effect in initiating and maintaining control loops organization-wide, but not enough for sure, since the number of studies reporting SPC success at maturity level 4 or 5 are still few.

Other challenges that might hinder SPC implementation for software include process related issues such as rational sampling and sub-grouping, measurement and analysis practices in use, effective metric selection, and the use of correct statistical methods (Figure 1.1). Contrary to most other process improvement difficulties that impediments change; these factors are technical, not managerial. The factors may serve as better starting points for cultural change in emergent software organizations that might have difficulty in corporate management but are generally more open to adopt technical solutions.

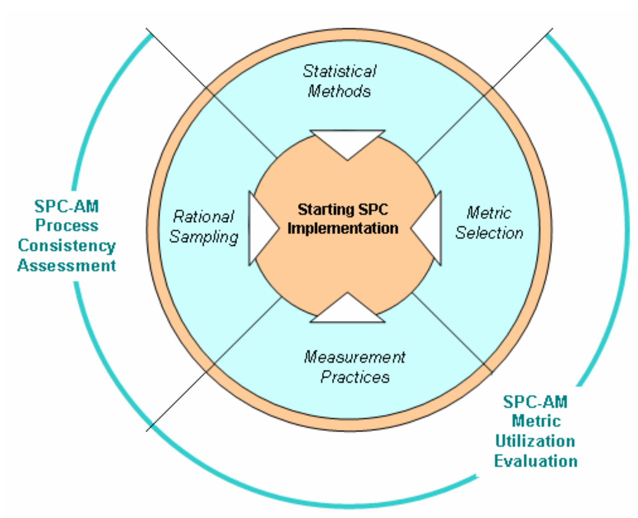


Figure 1.1 Factors That Contribute To Starting SPC Implementation

Accordingly, while developing the Assessment Model for Statistical Process Control (SPC-AM), we considered two basic requirements for SPC implementation, and focused on finding ways to resolve the difficulties brought by these requirements for an emergent organization: Rational sampling of process executions and data, and metric data utilization (or suitability) for statistical analysis.

The purpose of rational sampling is to obtain and use data that are representative of the performance of the process with respect to the issues being studied. If we can consider that observations are made under essentially the same conditions and that differences between the measurements are primarily due to common cause variation, then we are very likely that we rationally group the observations [35]. Since we want to sample process executions as being from a single and constant system of chance causes, we developed a clustering method based on the idea of process consistency assessment. We recommend describing each process execution in a number of process attributes such as inputs, outputs, activities, roles, and tools and techniques. Process consistency is assessed on a matrix for similarity in process attribute values of process executions. If repetitions of a process show similarity in terms of these attributes, then we assume that we rationally group the executions, each group being consistently performed.

The second requirement is metric utilization. This includes elaboration of basic measurement practices as well as metric data existence and characteristics. Measurement practices should be performed for a specific purpose [28][29][30] and metrics should be uniquely understood (e.g. by providing operational definitions) to enable consistent implementation. Operational definitions tell people how measurements are made so that others will get the same results if they follow the same procedures. To evaluate metric utilization, we identified a number of metric usability attributes that are metric identity, data existence, data verifiability, data dependability, data normalizability, and data integrability. We developed questionnaires based on these attributes for base and derived metrics separately. Questionnaires include a rating system based on answers to questions, and accordingly, evaluate the usability of a specific metric for applying SPC.

To refine and validate our model, we implemented three case studies in multiple-case-study context. We worked on task management, review, test development processes and related metrics of different organizations. The first case investigated utilization of estimation capability and effort variance metrics of task management process of a project-based working software organization. In the second case, we worked on non-conformance detection efficiency, non-conformance resolution efficiency, review open period, and review open period with respect to non-conformances metrics of review process of a system and software development organization targeting to achieve CMMI L3. In the third case, we worked on test design, test procedure development, and test development peer review processes of a system and software development organization having SW-CMM L3. Although the works of these processes can be considered as separate case studies, we evaluated the results considering their inter-relations. We investigated the utilization of productivity and percent of internal review effort metrics for test design and test procedure development processes, and the utilizations of action item density, action item detection efficiency, and action item resolution efficiency metrics for test development peer review process. As control chart is one of the most sophisticated data analysis tools within SPC, we demonstrated practical evidence on the utilization of SPC via control charts.

Multiple case study results showed us that with established guidelines for rational sampling and metric utilization, emergent organizations can apply SPC techniques and attain the ability to understand its processes based on quantitative data.

1.3. Roadmap

In Chapter 2, we establish an overall understanding about Statistical Process Control. We explain the fundamental concepts like process stability and capability. We describe the tools used to support SPC, and give details on control charts as a sophisticated SPC tool. We provide a survey of the literature on SPC implementations for software as well as on measurement theory and practices.

In Chapter 3, we provide the details related to the assessment model and the assessment process. We describe basic components of the model and explain the assets developed for use in the assessment.

Chapter 4 is devoted to the refinement and validation of the assessment model. We describe the approach and the work plan. We provide design principles of the case studies, and explain the criteria for selecting the cases. We finally describe case study characteristics, and provide details related to each case study implementation.

In Chapter 5, we discuss multiple case study results and summarize lessons learned during case study implementations. We finally provide our conclusions on the model and its implementations, and discuss the contribution of the work that we performed. We also state candidate subjects for future work.

CHAPTER 2

BACKGROUND

The principles of SPC strongly support process management which we need to produce high-quality and on-time products meeting (and even exceeding) internal and external customers' expectations. Process management deals with producing high quality products by focusing on the processes that are used in production. If we speak for software engineering, it is about successfully managing the processes associated with developing, maintaining, and supporting software products and software-intensive systems [35]. If we can control (and predict the behavior of) the processes applied in producing software, then we are very likely to better plan the performance of these processes, monitor their progress, and take corrective actions in case of discrepancies. And if we can perform these practices organization-wide and for all our software projects under all cases, then we have the chance of producing high-quality and on-time products meeting customers' expectations. In other words, a predictable process is attained by applying SPC techniques.

2.1. SPC Concepts and Tools

The principles of statistical process control state that by establishing and sustaining stable levels of variability, processes will yield predictable results [5]. We can then say that the processes are *under statistical control*. Controlled processes are stable processes, and stable processes enable us to predict results as basis for planning, monitoring, and improving.

Variability in process behavior is observed by defining and monitoring several attributes or variables representing the outcomes of the process. The number of defects found during unit testing, the number of requirements that are changed after requirements analysis phase, etc. may be used to understand the behavior of the processes they represent. We call this behavior as *process performance*. It is a state to understand how the process is executed, upon which we can make evaluations to direct process improvement.

Variation exists in all data and consists of both noise (random variation) and signal (non-random variation). The values must be filtered somehow to separate the signals from the noise that accompanies them, since acting on noise as if it were signal may increase the variability in process results. This filtering may be based subjectively upon a person's experience and assumptions, or it may be objectively based on a more formalized approach. SPC and its associated control charts developed by Shewhart in 1920s serve as the most-widely used formalized approach to handle the variation in a process.

When all signals have been removed and prevented from recurring in the future, then we have a *stable* process. We have a single and constant system of chance causes, and we can confidently predict results. Having a stable process, however, does not mean that process performance is satisfactory; the process must also be *capable*. If variations in the characteristics of the product and in the operational performance of the process, when measured over time, fall within the ranges required for business success, then we have a capable process. Understanding the capability of the sub-processes that make up each software process is the first step in making progress towards quantitative process improvement [19].

The aim of statistical process control is first to detect non-random variation (signals) in the process as basis for providing process control; and second to demonstrate the random variation (noises) in the process (already under statistical control) as basis for monitoring and improvement. SPC tools including control charts are described in subsequent sections.

2.1.1 Tools Used for SPC

The basic tools used for statistical control are summarized below [36]:

- Check Sheet: Check sheets are good means for collecting data efficiently, reliably and easily. As the detail and characteristics of data are different, check sheets are designed specifically considering the particular needs. Metric datasheets are used extensively in order to represent the data in the desired format.

- Cause-and-Effect Diagram: Cause-and-effect diagrams are useful tools to visualize, categorize and rank potential causes of a problem, a situation or any outcome. They are also named as fishbone diagrams because of their shapes and are usually formed as a result of a discussion or a brainstorming session of a group of people.

- Scatter Diagram: In a scatter diagram, data for two variables are collected in pairs (x_i, y_i), and each point y_i is plotted against corresponding x_i. This is a useful plot for identifying a potential relationship between two process characteristics. Scatter diagrams may be used for regression analysis.

- Run Chart: Run charts are specialized, time-sequenced form of scatter diagrams that can be used to examine data quickly and informally for trends or other patterns that occur over time. They dynamically observe performance of one or more processes over time. They are useful for visualizing performance after a process change.

- Histogram: Histograms show the frequency distribution of data in a sample. The first step to draw a histogram is to categorize the data into classes with equal ranges. Then the number of data in each class is found and depicted with bars on the graph. The data represents the state of a system at a certain time; thus there is no time dimension. Histograms are quite practical to visualize central tendency and skewness of an attribute.

- Bar Chart: Bar charts are like histograms. But they are not only used for depicting the frequencies of occurrences, but also for showing any numerical value of the attribute.

- Pareto Chart: Pareto chart is another form of bar chart. However, the occurrences are ordered with respect to their frequencies. Pareto charts are good means to visualize the ranking of an attribute among different categories.
- Control Chart: Control charts are sophisticated statistical analysis tools, which include upper and lower limits to detect any outliers. They look like run charts, but with the control limits and center line. They are frequently used in SPC analyses and described in detail in the following section.

2.1.2 Shewhart's Control Charts

In 1920s Shewhart was working on the concept of quality control and brought the idea that each process is driven by forces of variation. Variation was resulting in loss of quality by causing inefficiency and waste. Shewhart categorized sources of variation into two [5]:

- Variation due to phenomena that is natural and inherent to the process and whose results are common to all measurements of a given attribute,
- Variations that have assignable causes that could have been prevented.

The concept is represented in equation form as follows:

[total variation] = [common cause variation] + [assignable cause variation]

Common cause variation is the variation in process performance due to normal or inherent interaction among the process components (people, machines, material, environment, and methods). Common cause variation of process performance is characterized by a stable and consistent pattern of measured values over time. Variations in process performance due to assignable causes, on the other hand, have marked impacts on product characteristics and other measures of process performance. Assignable cause variations arise from events that are not part of the normal process. They represent sudden or persistent abnormal changes to one or more of the process components. These changes can be in things such as inputs to the process, the environment, process steps themselves, or the way in which the process steps are executed.

During his studies at Bell Labs in 1920s, Shewhart proposed that it is possible to define limits within which the results of routine efforts lie to be economical. Variation in the process outcomes resulting in values out of these limits indicated that the process is not performed economically. To detect assignable causes, Shewhart utilized statistics and control charts, foundations of which are listed below [37]:

1. Shewhart's charts always use control limits which are set a distance of three sigma-units on either side of the central line.
2. In computing three-sigma control limits one must always use average dispersion statistic or a median dispersion statistic.
3. The conceptual foundation of Shewhart's control charts is the notion of rational sampling and rational subgrouping.
4. Control charts are effective only to the extent that the organization can use, in an effective manner, the knowledge gained from the charts.

Shewhart control chart model depends on hypothesis testing. After a sample of data (sufficient enough to represent the whole) is collected, its mean and variance are calculated. Individual data values are depicted as points in a time series graph with respect to control limits (Figure 2.1). Centerline is the mean, and lower and upper control limits (LCL and UCL) are derived from the mean and variance by the formula "Mean ± 3 Standard Deviation". Control chart defines the voice of the process since it is the data itself that determines the limits. Data values are analyzed with respect to upper and lower control limits together with their location in the chart. As long as the process values represent the chance causes, the process shows controlled variation and is under control. However, any single value representing an assignable cause indicates that the process is out of control, and an investigation of the reason for the assignable cause is required. Then necessary actions are taken and measurements are repeated. The charts are redrawn with existing data values, and this process is repeated until no evidence remains for the existence of assignable causes. Once the process is brought under control, further improvement activities are implemented to minimize the effect of chance causes.

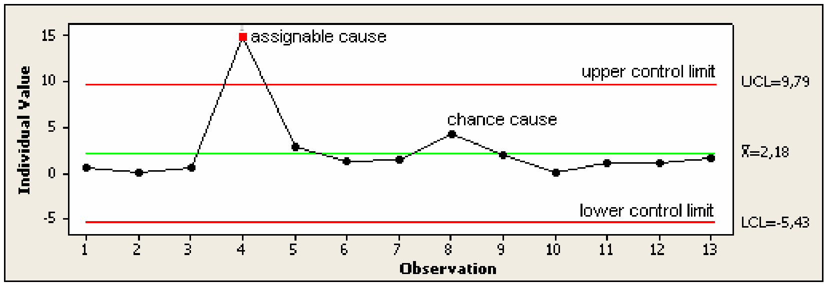


Figure 2.1 Example Control Chart

Control charts are not used only for determining process predictability. They provide a means to listen to the voice of the process, and to identify signals that have the potential for process improvement. From this perspective, control chart is an important tool in the process improvement toolkit. It allows to create a target where the current process is operating as consistently as possible, to drive the process toward that target, and to judge whether the process has come close to the target in practice [37].

The measurement can be performed by means of either variables or attributes. Burr and Owen define a variable as "*measure of a product that can have any value between the limits of the measurement*", while an attribute as "*count of things which may or may not be present in the product*" [38]. Variables data are usually measurements of continuous phenomena. Elapsed time, effort expended, years of experience, cost of rework, and memory utilization are examples for variables data. Attributes data, on the other hand, occur when information is recorded only about if an item conforms or fails to conform to a specified criterion or set of criteria, and almost always originate as counts. The number of defects found, the number of defective items found, the number of source statements of a given type, the number of people with certain skills or experience, the percent of projects using formal code inspections are examples for attributes data.

The nature of these two measurement categories necessitates different statistical analyses. Control limits for attributes data are often computed in ways different from control limits for variables data. Whether the data should be treated as attributes or variables type gains importance here. The key to classifying data as attributes data or variables data depends closely on how the data are collected and used, rather than on whether they are discrete or continuous [35]. For example, the number of working days in a month might be viewed as attributes data if used as a numerator to compute the proportion of a month available for working (e.g. 20 working days per 30 days in April), or as variables data if used as a denominator to normalize some other measure of activity (e.g., number of design documents produced per month – per 20 working days in April, per 23 working days in May, etc.). Unless we have a clear understanding of the distinctions between the two kinds of data, we can easily fall victim to inappropriate control-charting methods.

There are several types of control charts. We can use different charts for subgroup averages, moving averages, and individual values [37]. It is recommended to use Xbar-R chart or Xbar-S chart for subgroups of, and X-chart or XmR chart for individuals of variables data. The types of control charts to use for attributes data, on the other hand, are p-, np-, c-, and u-charts as well as XmR chart for counts or rates. Below are further explanations on these control charts [35].

Xbar-R Chart: Averages and range chart is used to portray process behavior when we collect multiple measurements within a short period of time under basically the same conditions. Measurements are then grouped into self-consistent sets (subgroups) that can reasonably be expected to contain only common cause variation. The results of the groupings are used to calculate process control limits.

Xbar (average) charts answer the questions as "what is the central tendency of the process?" and "how much variation has occurred from subgroup to subgroup over time?". The corresponding R (range) charts indicate the variation (dispersion) within the subgroups. It is advised that range charts be used only when there are 10 or less observations in each subgroup.

Xbar-S Chart: Averages and standard deviation chart is used instead of Xbar-R charts when subgroup size is larger than 10. S charts based on averages of the standard deviation within subgroups give tighter control limits, which brings increased sensitivity to assignable causes. As the size of the subgroup increases, it becomes increasingly difficult to ensure homogeneity of the subgroup. Therefore, for reliability, selection of the subgroup size should be dictated first by the homogeneity of the subgroup and second by the subgroup size.

X-Chart: When measurements are spaced widely in time or when measurement is used by itself to evaluate or control a process, a time-sequenced plot of individual values, rather than averages, appears. This means that the subgroup size is 1.

An individual plot can detect more readily the following conditions than an Xbar-R chart: cycles (regular repetitions of patterns), trends (continuous movement up or down), mixtures (presence of more than one distribution), grouping or bunching (measurements clustering in spots), and relations between the general pattern of grouping and a specification.

XmR Chart: Individuals chart is frequently complemented by a corresponding moving range chart which depicts successive two-point moving ranges. This combination of charts for individual observations and moving ranges is called and XmR chart. XmR charts are especially useful to view trends in the process.

The idea behind XmR chart is that, when subgroups can easily include nonrandom components, we minimize the influence that nonrandom effects have upon estimates for sigma by keeping the subgroups as small as possible. The smallest possible subgroup size is 1. There is no way to estimate sigma from a single measurement so that we do the next best thing: We attribute the changes that occur between successive values to the inherent variability in the process. The absolute values of these changes are called two-point moving ranges.

When median moving range is used instead of the average moving range to compute the limits for an XmR chart, then we have “X and median mR” chart. The median moving range is frequently more sensitive to assigned causes when the moving range contains several very large values relative to the rest of the moving range values. Several high range values unduly inflate the average moving range and cause the upper and lower limits to expand.

np-Chart: An np-chart is used when the count data are binomially distributed and all samples have equal areas of opportunity. For example, when there is 100%

inspection of lots of size n (n constant) and the number of defective units in each lot is recorded.

p-Chart: A p-chart is used instead of an np-chart when the data are binomially distributed but the areas of opportunity vary from sample to sample. A p-chart is appropriate in the inspection example given for np-chart, if the lot size n were to change from lot to lot.

c-Chart: A c-chart is used when count data are samples from Poisson distribution and the samples have equal-sized areas of opportunity. C-charts are suggested, for example, when tracking the number of defects found in lengths, areas, or volumes of fixed (constant) size.

u-Chart: A u-chart is used instead of a c-chart when the count data are samples from a Poisson distribution and the areas of opportunity are not constant. Here, the counts are divided by the respective areas of opportunity to convert them to rates. A u-chart is more flexible than a c-chart because the normalizations that it employs enable it to be used when the areas of opportunity are not constant.

An XmR chart can be used in any of the above situations described for attributes data as well as when neither a Poisson nor a binomial model fits the underlying phenomena or when little is known about the underlying distribution. However, an XmR chart is not a reasonable choice when the events are so rare that the counts are small and values of zero are common (then the discreteness of the counts can affect the reliability of the control limits). If the average of the counts exceeds 1.00, an XmR chart offers a feasible alternative to the traditional attributes charts.

Wheeler suggests the following tests for detecting the assignable causes in a control chart [37] ("sigma" means standard deviation):

- Test-1: A single point falls outside the 3-sigma control limits.
- Test-2: At least two out of three successive values fall on the same side of, and more than two sigma units away from, the centerline.
- Test-3: At least four out of five successive values fall on the same side of, and more than one sigma unit away from, the centerline.

- Test-4: At least eight successive values fall on the same side of the centerline.

Tests 2, 3, and 4 are called *run tests* and are based on the presumptions that the distribution of the inherent, natural variation is symmetric about the mean; that the data are plotted in time sequence; and that successive observed values are statistically independent. The symmetry requirement means that the tests are designed primarily for use with X-bar and individuals charts. Strictly speaking, they are not applicable to R charts, S charts, or moving range charts [35]. Using test 1 avoids the need to make assumptions about the distribution of the underlying natural variation.

As Wheeler points out, these four tests are conservative and practical subset of the much larger body of run tests that have been used from time to time in industrial settings. Each additional run test increases our chances of detecting and out-of-control condition; however, it also increases our chances of getting a false alarm. Here the important point is that the decision to use a test should be given before looking at the data. Determining the frequency with which a specific test leads to false alarms would be wise to identify its effectiveness.

2.2. Literature on SPC for Software

2.2.1 Implementations of SPC for Software

Analyzing a Mature Software Inspection Process Using SPC

This is a presentation that explains the results from a cooperative effort where Software Engineering Institute and the Space Shuttle Onboard Software Project experiment applying SPC analysis to inspection activities [10]. During the study; project process descriptions are reviewed, data definitions are verified and validated, and experimentation and analysis are conducted. Since SPC analysis assumes data come from different sources, six functional areas of the project are treated separately. Within each functional area, inspection process data is graphed as a function of calendar time over four releases, and each plot is examined for process stability. Correlation analysis is conducted to determine validity of cause-effect relationships, and process performance is analyzed for each functional area.

Control charts are depicted and examined for the following metrics in search of stability: Preparation hours per inspection, preparation hours per modified SLOC, error per preparation hours, error per modified SLOC, preparation hours per inspector, preparation hours per inspector per modified SLOC, inspectors per inspection, and SLOC per preparation hours per inspector. Initial observations show that control charts dealing with preparation hours and/or modified SLOC appear to exhibit bunching or mixtures implying two or more processes. The charts for SLOC/preparation time per inspector and SLOC/preparation time per primary inspector metrics are observed as stable. Further observations from the study are given below:

- At least two varying executions of the code inspection process appear to occur based on the amount of design or code material being inspected.
- The characteristics of data only inspections are different from code inspections and require separate analysis. Also, the characteristics of code re-inspections are different from initial inspections and require separate analysis.
- Code inspections of greater SLOC require more time; but the rate of review of the material is faster. Code inspections of greater SLOC have more errors and code inspections of lesser SLOC (i.e., less than 50) have infrequent errors.

After the study, the presenters state that examining, normalizing, and determining stable process performance variables takes considerable effort; consistency in data collection and reporting is imperative; and clarifying and understanding how the data is defined is crucial to knowing what the data represents.

Statistical Process Control: Analyzing a Space Shuttle Onboard Software Process

Florac and his friends describe the experiences of a study on the application of SPC based on the data and analysis from a collaborative effort between the Software Engineering Institute and the Space Shuttle Onboard Software Project [11]. The study and its results are explained above [10], and the lessons learned from the study together with the suggestions are the subject of this article.

The authors state the following notions as important to consider for fully appreciating how control charts are used to measure and analyze software processes. They claim that awareness and attention to these factors and others are critical for successful use of control charts to analyze software processes: Selecting key and critical processes, providing operational definitions, addressing issues of data homogeneity and rational sub-grouping, using the correct control charts, understanding multiple-cause systems and mixed caused systems, finding and testing trial limits, and recalculating limits.

These factors were applied for the Space Shuttle Onboard Software Project. The authors conclude that effective use of SPC requires a detailed understanding of processes and willingness to pursue exploratory analyses. Coordination Manager of the project explained that applying SPC to their software development activities helped in the following ways:

- Fully understanding process behaviour provided an understanding of the reliability of human processes.
- Understanding the inherent process variation established pragmatic bounds on management expectations (e.g., distinguishing variations due to people problems from variations that are other process problems).
- Understanding patterns and causes of variation let the Space Shuttle Onboard Software Project understand the dynamics affecting process behaviour and what “stable” meant in a particular environment.

Practical Applications of Statistical Process Control

Weller provides a distinct case in his article by presenting details on SPC implementation to analyze inspection and test data in a software organization [12]. He proposes that in order to regard defect density as an indicator of product quality, he first wants to be sure that inspection process is stable in the organization. He uses X and moving range charts for the lines of code inspected per hour for each inspection, and achieves a stable inspection process after removing the outliers from the dataset. Then he draws u-chart for the defect density data for each inspection. By these findings, he makes reliable estimations for inspection effectiveness and gains an insight on when to stop testing.

The results of the analysis are discussed with the project teams at their weekly meetings, for three main reasons: It sends a message that the data is being used to make decisions on the projects; keeping the estimates and data in front of the teams make them aware of the progress toward the quality targets; and they want to avoid the problem of "metrics are going into a black hole" which causes metric programs to fail.

Weller states that they gained a fact-based understanding of many of their release processes. They were able to set quality goals, measure the results, and predict a post-ship rate with confidence. The cost for this benefit included analysis of inspection data, collection of unit test data, and analysis of integration and system test data. He concludes that SPC implementation helped to understand and predict the release quality, and the development process controlling that quality.

Quantitative Quality Management through Defect Prediction and SPC

This article [13] describes the approach of quantitative quality management through defect prediction and statistical process control that is employed at Infosys, a large ISO-certified software house that has been assessed to be at level 5 of the CMM. In this approach, a quality goal is set for a project in terms of the defect density delivered. To achieve the goal, the defect levels for different phases in the process are estimated using past data. During the execution of the project, the actual defect numbers are compared with the estimates to see if the project is progressing satisfactorily towards achieving the goal, or some correction is needed. To further improve the control and provide early warnings, the phase-wise control is complemented with activity-level control using statistical process control.

For reviews, based on past data, control charts are built for different types of reviews. From the control charts, the review capability baseline is established, which gives the control limits for key parameters like defect density found in reviews, preparation rate, review rate, etc. These limits are used to evaluate a review. At the end of each review, it is checked if the defect rate is within the limits. If yes, nothing needs to be done. If no, then other rates are checked and based on evaluation, some corrective and preventive actions may be taken. Guidelines are provided for evaluation.

For unit testing, control charts are built, mostly for defect density. The results of a unit testing are checked against the control limits. Again, action may be taken if the results are out of the limit, and guidelines have been provided for evaluation

Statistical Process Control to Improve Coding and Code Review

The article explains details related to SPC implementation via control charts to control variation in the coding and code review processes [14]. The authors use data from process automation and consumer electronics projects developed in C++. Project sizes range from 150 to 400 function points. In the coding-and-code-review scenario, they plot the values across classes or files, and group projects into different categories, such as process automation, consumer electronics, drivers, Web-based software, and embedded systems. They study the charts for preparation speed, review speed, defect density for code review, and defect density for testing; and plot the charts in the order of the units coded. The code review findings will serve as in-process feedback to the coding process, establishing a closed loop for continuous process improvement within the project.

The process is found as stable with respect to preparation speed, review speed, defect density for code review, and defect density for testing metrics with several assignable causes. Further observations from the study are given below:

- When studied together, the charts for preparation speed, review speed, and defect density for code review give further insight into each unit's quality and the review process's effectiveness in terms of effort expended for the review.
- When studied together, defect density for code review and defect density for testing charts give better insight into each unit's quality and the review's effectiveness in terms of defect detection.

The authors claim that the key to successful chart analysis lies in recognizing the indications the chart provides, being able to map them to the change that has occurred in the process, and using the information to continuously improve the process. They also note that the interpretation of the chart and identification of corrective actions depends on the process being analyzed, the project type, and the team's expertise and experience. So, the team members involved in the process should perform the analysis.

Utilization of Defect Density Metric for SPC Analysis

This is one of the unique articles that elaborate SPC implementation for software processes. The article [15] describes the difficulties and suggestions in application of SPC to a CMM Level 3 organization using defect density metric. It is a part of a broader study completed within a master thesis [39]. It discusses the defect density metric and demonstrates that the metric requires a precise definition of defect as well as products size for different phases of software development to be used for statistical process control. The authors prefer to use XmR charts for tracking defect density instead of the popular u-chart, and show that XmR chart is more appropriate for analyzing defect density data.

After summarizing problems and solutions on defect density metric, the authors provide implementation details. They analyze defect density while performing research studies on usability of SPC techniques. In the company, the data of all defects found during a review, test, or audit have been collected and tracked through Problem Reports (for code defects) and Document Change Requests (for document defects) since the foundation of the company in 1998. Each defect on a trouble report is given a priority, which is classified as low, medium, high, very high and other. However, after collecting the data, they combine 5 priority categories within 3 groups: 1) Combining high and very high, 2) For medium, and 3) Combining low and other. This categorization is made by the assumption that similar attention can be paid to the defects in priorities low and other.

The code size is collected for each CSCI in terms of Source Lines of Code, excluding comment lines and blanks. Considering process control purposes, the authors decide to restrict the analysis to requirements and design documents and define size measures as follows: Requirements documents (SRS and IRS) – The number of requirements is used to compute size; Design Documents (SDD and IDD) – The number of pages is used to compute size. The authors also compute the cumulative number of defects for each document. As the size of a document remains almost the same throughout the project, the defect density value gradually increases as more defects are detected. Therefore, the authors restrict the analysis for design and maintenance phases.

The document size is gathered for each version and the size of the last version which is already released at the end of project phase is used for defect density measurement. Afterwards, XmR charts are drawn for each project phase-priority-document type combination. The observations are in the order of the document preparation times from past to the future. It is observed that the process is under control and all the variation comes from inherent process characteristics. The authors prepare similar charts for design and maintenance phases with different priorities; however, they can not obtain high effectiveness.

Based on the observations the authors claim that it is necessary to make a precise definition of defect and categorize defect data so that the data becomes meaningful for SPC analysis. Moreover, the size measure should be distinct and well-defined for different work products. While computing size, it is also important to obtain separate measures for different sections of the work products.

Statistical Process Control Applied to Software Requirements Specification Process

This presentation [40] explains the experience of MITRE Corporation in a government agency that reverse-engineer the existing software requirements while re-developing legacy systems. Five teams are assigned to reverse engineer related sets of functional requirements, and the author is assigned as a consultant to support the agency in the proper specification of the requirements.

The presentation includes a number of examples that illustrate the application of control charts applied to the requirements specification process. The examples show some requirements as initially specified by the teams and followed by the authors critique against the critical attributes of requirements. Each violation against the critical attributes is recorded as a defect to be used to construct control charts. Below are the summaries for some examples:

- One example shows a control chart of all teams' attempts at the initially specification of the requirements. This is before they received guidance on the critical attributes (that is, they are not yet following a consistent process). The control chart showed that the process is immature and out of statistical process control.
- Another example shows a control chart of all teams' subsequent attempts at the specification of the requirements. New sets of requirements are included. The teams are trained in the critical attributes and most resolve the critique issues. An anomaly occurs with the second team's effort, and causal analysis reveals that the second team does not implement the critique's findings nor analyze new requirements against the critical attributes.
- Yet another example shows a control chart of all teams' subsequent attempts at the specification of the requirements. New sets of requirements are included. Management ensures that the second team resolves the issues identified in the critique and that they analyze additional requirements against the critical attributes. The requirements specification process is under statistical process control.

The examples demonstrate the use of SPC applied to the requirements specification process. Many control charts are constructed and analyzed. The author claims that the ones explained here are selected to succinctly demonstrate their use. He notes that the use of statistics using SPC control charts and other statistical methods can easily and effectively be used in a software setting. SPC can identify undesirable trends and can point out fixable problems and potential process improvements and technology enhancements. Finally, the author argues that using SPC, beginning with requirements analysis, can provide the biggest payoff. It is a well-known fact that if requirements are properly defined early in the development life cycle, the migration of problems into the later phases will be mitigated.

2.2.2 Guidance on SPC for Software

Application of Statistical Process Control to Software Processes

Lantzy is one of primary authors that mention the application of SPC concepts for software. In his article [8], he summarizes the concept of SPC, gives some practical examples from manufacturing industry, and offers a set of transformations on SPC principles for use in software engineering. He argues that the transformation of user requirements into software is dominated by cognitive activities and higher level cognition increase variances in productivity and quality, making the application of SPC more difficult. Lantzy states that effective SPC application depends on the ability of managers to negotiate a prioritized list of quality characteristics and acceptable tolerances with their customers and to apply SPC principles in a manner that assures conformance of the software product to that prioritized list. With this statement, he implies that the process should be designed based on the product goals. The tailoring of the process includes the tailoring of process metrics to the quality characteristics of the end-product.

Lantzy outlines a seven-step guideline for successful application of SPC principles to the software process: Negotiate a set of prioritized software quality characteristics with the customer. Design, specify, and implement a software process capable of producing the desired software product. Establish process owners and empower them. Establish metrics for processes that correlate to the quality characteristics established for the end-item software product. Employ control charting or comparable techniques to determine the stability of each process. Bring processes in control by eliminating all special causes of variation. Continuously improve processes in order to bring control limits within tolerances so that the end-item software product meets customer requirements.

After summarizing a case-study (POST project, U.S. Navy) validating the guideline described above, Lantzy concludes that SPC is not just a measurement discipline, but also a product planning and assurance philosophy that recognizes the variation inherent in all processes.

Statistical Process Control for Software?

Card discusses the utilization of SPC for software in his article [9] by mentioning some objections and possible implementation problems. He states that, as one objection, software development process does not involve repeated delivery of equivalent services or the fabrication of identical products. Another objection is the lack of a perfect measure of the attributes, which actually underlies the importance of metric definition. However, he argues that SPC does not rely on having a perfect measure, since SPC analysis is meant only to give some insight into how the process is functioning and it does not have to provide total visibility. He recommends beginning with a model of the process and then selecting techniques to monitor performance, in implementing SPC. He provides an example of a control chart to track testing efficiency, related to his approach.

Card outlines possible implementation problems under six issues: undefined process, poor choice of measures, focus on individual or small events, incorrect computation of control limits, failure to investigate and act, and lack of training. He concludes that SPC principles can be beneficial for a software organization although formal statistical control techniques may not be used.

Practical Software Measurement: Measuring for Process Management and Improvement

This is a guidebook that explains the perspectives of process measurement and elaborates the requirements of process management based on measurement practices [19].

The concept of process management is founded on the principles of statistical process control. These principles hold that by establishing and sustaining stable levels of variability, processes will yield predictable results. Predictable results should not be construed to mean identical results. Results always vary; but when a process is under statistical control, they will vary within predictable limits. If the results of a process vary unexpectedly—whether randomly or systematically—the process is not under control, and some of the observed results will have assignable causes. These causes must be identified and corrected before stability and predictability can be achieved.

Controlled processes are stable processes, and stable processes enable you to predict results. This in turn enables you to prepare achievable plans, meet cost estimates and scheduling commitments, and deliver required product functionality and quality with acceptable and reasonable consistency. If a controlled process is not capable of meeting customer requirements or other business objectives, the process must be improved or re-targeted.

At the individual level then, the objective of software process management is to ensure that the processes you operate or supervise are predictable, meet customer needs, and (where appropriate) are continually being improved. From the larger, organizational perspective, the objective of process management is to ensure that the same holds true for every process within the organization.

There are four responsibilities that are central to process management:

- Define the process,
- Measure the process,
- Control the process (ensure variability is stable so that results are predictable),
- Improve the process.

There are five perspectives that are central to process measurement: Performance, stability, compliance, capability, and improvement and investment.

Performance: What is the process producing now with respect to measurable attributes of quality, quantity, cost, and time?

The first step in controlling a process is to find out what the process is doing now. All processes are designed to produce results. The products and services they deliver and the ways they deliver them have measurable attributes that can be observed to describe the quality, quantity, cost, and timeliness of the results produced. If we know the current values of these attributes, and if a process is not delivering the qualities we desire, we will have reference points to start from when introducing and validating process adjustments and improvements.

So the first concern when measuring for process management and improvement is to understand the existing performance of the processes we use—

what are they producing now? Knowing how a process is performing will enable us to assess the repeatability of the process and whether or not it is meeting its internal and external needs (Notice the word "how" rather than "how well"). When measuring process performance, the purpose is not to be judgmental, but simply to get the facts. Once the facts are in hand and we know the current levels and variabilities of the values that are measured, we can proceed to evaluating the information from other perspectives.

Stability: Is the process that we are managing behaving predictably?

Measures of process performance quantify and make visible the ability of a process to deliver products with the qualities, timeliness, and costs that customers and businesses require. When measurements of process performance vary erratically and unpredictably over time, the process is not in control. To attain control, we must ensure first that we have a process whose variability is stable, for without stability we cannot predict results. So another important property associated with any process is that of process stability.

How do we know if a process is stable? We must first define what we mean by stable, and then we must find ways of measuring appropriate process and product attributes to determine if stability has been achieved. If process performance is erratic and unpredictable, we must take action to stabilize that process.

Compliance: Are the processes sufficiently supported? Are they faithfully executed? Is the organization fit to execute the process?

Stability of a process depends on support for and faithful operation of the process. Is the process supported such that it will be stable if operated according to the definition? Is the process, as defined, being executed faithfully? Is the organization fit to execute the process? Questions of this sort address the issue of process compliance.

Capability: Is the process capable of delivering products that meet requirements? Does the performance of the process meet the business needs of the organization?

Having a stable and compliant process does not mean that process performance is satisfactory. The process must also be capable. Capable means that variations in the characteristics of the product and in the operational performance of the process, when measured over time, fall within the ranges required for business success. Measures of process capability relate the performance of the process to the specifications that the product or process must satisfy.

Improvement: What can we do to improve the performance of the process? What would enable us to reduce variability? What would let us move the mean to a more profitable level? How do we know that the changes we have introduced are working?

If a software process is not capable of consistently meeting product requirements and business needs, or if an organization is to satisfy ever-increasing demands for higher quality, robustness, complexity, and market responsiveness while moving to new technologies and improving its competitive position, people in the organization will be faced with the need to continually improve process performance. Understanding the capability of the sub-processes that make up each software process is the first step in making progress towards process improvement.

Statistical Process Control for Software Projects

Radice provides a tutorial including information on concepts, use, and techniques of SPC, together with practical experiences [21]. He gives various definitions of SPC from the literature, and discusses on which processes to apply SPC as well as pre-conditions for SPC. He explains data characteristics and causes of variation, and states not all SPC techniques are applicable for software processes. Radice argues that SPC can be started at CMM Level-1, if there is consistent process execution and sufficient data. However, one should determine which processes are significant business drivers before applying SPC.

Radice states that the following software processes might be considered for SPC:

- Life cycle step processes (e.g., requirements analysis, design, code, test, maintenance);
- Recurring processes (e.g., configuration management, training, planning, estimating, tracking, defect prevention, inspection, hardware utilization).

Radice recommends considering a number of pre-requisites for any process before applying SPC: The process has characteristics that contribute to significant business drivers (e.g., cost, quality, time, customer satisfaction). Process is defined and measurable, and performed with consistency and within a reasonable bandwidth. Measures are defined, and sufficient data points are available. Resultant data are accurate and have integrity within the selected process (e.g., reliable, stable over time, comparable). The process can be modified based on improvement analysis and feedback. Customer defined limits are available, or natural limits are known.

Measuring the Software Process: SPC for Software Process Improvement

This is a book [35] that explains specifically how quality characteristics of software products and processes can be quantified, plotted, and analyzed so the performance of software development activities can be predicted, controlled, and guided to achieve both business and technical goals. The book is an extension and elaboration of the guidebook "Practical Software Measurement: Measuring for Process Management and Improvement" [19].

This book is organized into eight chapters. The focus of Chapter 1 is to introduce the primary concepts associated with managing, measuring, controlling, and improving software processes. The motivation for using statistical process control is also discussed (that is, utilizing control charts for making process decisions and for predicting process behavior). The chapter begins by characterizing the term software process, especially as it is used in SPC applications. Issues of process performance, stability, compliance, capability, and improvement are briefly introduced (and elaborated throughout the book) since these form the basis for improving process performance. A section on measuring process behavior then follows. A framework for measuring process behavior is presented next and serves as the guiding structure for the rest of the book. The remaining chapters follow this framework with more detailed discussions, expanding on the activities associated with using statistical process control techniques for improving the software process.

The focus of Chapter 2 is to discuss the activities associated with measuring the software process. They include identifying process management issues, selecting and defining the measures, and integrating the measurement activities with the organization's processes. The idea here is to understand what you want to measure and why and to select appropriate measures that will provide insight into your issues. In Chapter 3, the specifics associated with collecting software process data are discussed. The principal tasks include designing methods and obtaining tools for data collection, training staff to execute the data collection procedures, and capturing and recording the data. Additionally, there is a discussion of many of the important tools available to analyze, understand, and explain causal relationships to the process performance data. In Chapter 4, the authors embark on the initial discussion of analyzing process behavior with Shewhart's control charts by graphically illustrating the concepts of process variation and stability. The basics of constructing control charts, calculating limits, and detecting anomalous process behavior are given to provide a basis for the ensuing chapters. Chapter 5 is dedicated to providing the information to construct and calculate limits for the several different control charts applicable to software processes. Examples of the calculations and charts are set in familiar software settings. Chapter 6 discusses a number of topics that arise when using control charts. Guidelines are offered for how much data is necessary for control charting, recognizing anomalous process behavior patterns, rational sub-grouping, aggregation of data, and insufficient data granularity. Chapter 7 provides insight on what actions to take after plotting data on process behavior charts. The actions involve removing assignable causes of instability, changing the process to make it more capable, or seeking ways to continually improve the process. The book concludes with Chapter 8. It provides ten steps for getting started using statistical process control, cites the experiences by some of those who have used statistical process control in a software environment, and addresses a number of frequently asked questions.

Can Statistical Process Control Be Usefully Applied to Software?

This is a presentation that discusses some of the pros and cons from industry use of SPC [17]. The focus in particular is on prerequisites for successful use of SPC and its business value. The presentation summarizes crucial points on SPC for software from the work of various authors including Keller, Meade, Burr, Hirsh, Heijstek, Wigle, Curtis, Card, and Barnard. These crucial points are gathered in the table below.

Table 2.1 Crucial Points on SPC for Software

Author	**Crucial Points**
Keller	SPC is important to managers Understand reliability and set expectations Fix the "right problem"
Meade	Plan → Informally stabilize → Stabilize → Establish capability Ensure management understands the intent of level 4 Understand your data Smaller programs are better able to use SPC (implies an emphasis on micro-level processes rather than macro-level (project-level))
Burr	Management use of SPC (data) at low maturity levels For change, set targets outside the 3 sigma limits Consistent (stable) process at team level Change of (management) culture
Hirsh	Tie improvement activities to business objectives (e.g., customer satisfaction) Sophisticated SPC charts are useless unless used Not everything you try will succeed Communication is important too Require SPC training for managers Poster boards outside offices as motivational tools
Heijstek	Importance of data quality for good analyses (Can you ever get enough data points from the same system in the dynamic telecoms environment?)
Wigle	We have to be prepared to answer the hard questions. Do we want "all" processes to be stable? (Does this contradict continual improvement? Is SPC impeding continual improvement?) Are we applying SPC at the decision-making level? What should be measured and statistically controlled?
Curtis	Statistics/measurement are important because of insight Individual differences can overwhelm every other factor There are other statistical techniques than control charts that add value
Card	Lack of well-defined business objectives is problem Need to probe nature of data and how it was collected Scholastic thinking implies need to think Lot of data problems when starting out (first emphasis is on stabilizing the process) Understand the data first, then try techniques
Barnard	Be sensitive to "mixing" of multiple (similar) processes (significantly different sizes of work product, data versus code inspections, inspections versus re-inspections) Watch for non-linear associations

Considering Statistical Process Control for Software

This is a tutorial that presents basic concepts of SPC for use in software industry [20]. The presentation covers the topics of CMM context for SPC, business context for SPC, statistical thinking, "informally stabilizing" the process, SPC techniques, and challenges to SPC for the software process. Below are the keynotes from each of these topics:

- *CMM context for SPC*: Process data is collected at the "process step" level for quantitative process management. Engineers use the data to drive technical decision making (e.g., design inspections, code inspections, test cases). Data collected at phase end or on monthly basis is too late for real-time control. High maturity organizations collect a lot of data at the sub-process level. To use data for control and comparison, data sources must be categorized by product family, application domain, etc. A few important business drivers determine the vital few measures (e.g., cost, schedule, quality). Level 4 of CMM emphasizes "quantitative management" rather than "statistical control". Levels 4 and 5 conceptually based on assignable and common causes of variation, and most level 4 and 5 organizations initially appraised using a "relaxed" interpretation of quantitative (statistical) management. Problems in reliably and consistently interpreting levels 4 and 5 are similar to the problems in interpreting levels 2 and 3 in 1990 – before the publication of Software CMM v1. For institutionalization, one should select "critical" processes to be quantitatively managed. A reasonable rule-of-thumb for institutionalization is that quantitative management has been in practice for 6-12 months. This depends on frequency of execution, and organizations go through an "informally stabilizing the process" phase. Organizations should demonstrate at least a pilot use of rigorous statistical techniques, such as control charts or prediction intervals.
- *Business context for SPC*: Related questions are "Is it possible to apply the concepts of statistical process control to the software process?", "Do we know how to measure software products and processes?", "Is a stable, predictable process meaningful in a rapidly changing, high-tech environment?", and

"Does a capable process really add business value in a world of "difficult" customers?". Managers have different decision making needs and time horizons than engineers. At the project and higher organizational levels, risk management in the face of uncertainty drives the decision making process. At the sub-process level, engineers take advantage of what is known about process performance, and know when new processes and technologies invalidate historical information.

- *Statistical thinking*: Statistics is the science of patterns in a variable world, and deals with the patterns of "chance". Statistics makes the invisible visible, including the invisible of what has not yet happened. Knowing what is possible with the current process may indicate the kind of management action necessary to achieve those targets. All work is a series of interconnected processes. All processes are variable, and understanding variation is the basis for management by fact and systematic improvement (understand the past, control the present, and predict the future – all quantitatively).
- *"Informally stabilizing" the process*: Arguably, only Level 3+ organizations have the consistently-performed processes necessary to consider SPC in a rigorous manner. Shewhart believed his work on operational definitions have been of greater importance than his development of the theory of variation and of the control chart. There are two criteria for operational definitions: Communication and repeatability. Poor operational definitions lead to process inconsistency and product variation that causes inconsistency. Inadequate contextual information leads to lack of traceability from data back to its original context. Data whose elements are combinations (mixtures or stratification) of values come from different sources such as variability of individuals and team composed of individual workers. Process capability may be determined for the organization, product line, project, team, and individual. The higher the level of analysis, the greater the variation, and the less useful the insight. High maturity organizations typically are doing systematic reuse with domain engineering and/or product lines/families.

- *SPC techniques*: Seven basic SPC tools include scatter diagrams, run charts, cause-and-effect diagrams, histograms, bar charts, pareto charts, and control charts. SPC implies control charts many times. Control charts let us know what our processes can do, so that we can set achievable goals. They represent the "voice of the process.", and provide the evidence of stability that justifies predicting process performance. For stability, the concern is "Is the process that we are managing behaving predictably?", and the business value is foundation for estimating (predicting) and making commitments. For capability, the concerns are "Is the process capable of delivering products that meet requirements?" and "Does the performance of the process meet the business needs of the organization?", and the business value is foundation for making commitments.
- *Challenges to SPC for the software process*: Myths about control charts include: 1) Data must be normally distributed before they can be placed on a control chart, 2) The control chart works because of the central limit theorem, 3) Observations must be independent, and 4) Data must be in control before one can plot them on a control chart. The challenges include too much variability, insufficient data, multiple and overlapping processes, confusing thresholds and control limits, incorrect statistical techniques, Hawthorne effect (measurement drives behavioral change), and causing dysfunctional behavior (motivational vs. information measurement), and management training.

2.3. Literature on Measurement Approaches and Models

Software Metrics: A Rigorous and Practical Approach

This is a book that covers the basics of measurement theory as well as the most known process, product, and resource measures [1]. Below are the keynotes as related to measurement theory.

A measure must specify the domain and the range as well as the rule for performing the measurement mapping. Both entity and attribute to measure should be explicit. Measures can be direct or indirect. Direct measures involve no other attribute or entity, and form the building blocks for assessment. Examples are size, duration, and number of defects. Indirect measures are derived from other measures. Examples include productivity, defect density, and efficiency.

Measurement mapping together with the empirical and numerical relation systems represent the measurement scale. Scales help us to understand which analyses are appropriate. Types of scales are nominal, ordinal, interval, ratio, and absolute, in increasing order of providing information. Nominal scale indicates a difference, just classification, and no ordering (e.g., flower names). Ordinal scale indicates the direction of the difference, and ranking with respect to ordering criteria (e.g., priority assignments). Interval scale indicates the amount of the difference, and differences of values are meaningful (e.g., calendar date). Ratio scale indicates an absolute zero, and ratios between values are meaningful (e.g., effort). Absolute scale indicates number of values (e.g., number of defects).

A mapping from one acceptable measure to another is called an admissible transformation. It is the transformations do not change the structure of the scale (e.g., feet mapped to inches). Understanding scale types enables us to determine when statements about measurements make sense. A statement involving measurement is meaningful if its truth value is invariant of transformations of allowable scales. The following statistical operations are allowed for each scale:

- Nominal: Mode, frequency.
- Ordinal: Median, percentile.
- Interval: Mean, standard deviation.
- Ratio: Geometric mean.

The Goal Question Metric (GQM) Approach

The Goal Question Metric (GQM) approach [28] proposes that measurement definition must be top-down as based on goals and models. The approach assumes that purposeful measurement is possible by specifying the goals for the organization and its projects, then by tracing those goals to the data that are intended to define those goals operationally and finally by providing a framework for interpreting the data with respect to the stated goals [29]. Result of the application of the GQM approach is the specification of a measurement system targeting a particular set of issues and a set of rules for the interpretation of the measurement data.

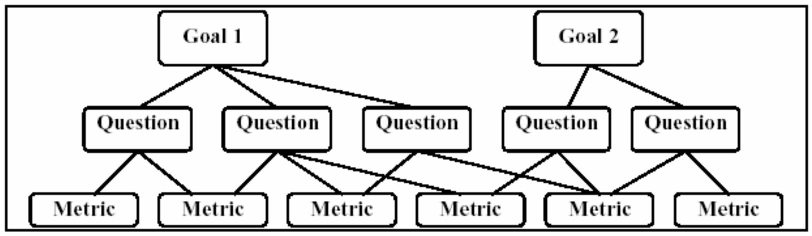


Figure 2.2 The Goal-Question-Metric Hierarchy

A GQM model is a hierarchical structure as shown in Figure 2.2. It starts with a goal specifying purpose of the measurement, object to be measured, issue to be measured, and viewpoint from which the measure is taken. Objects of measurement include products, processes, and resources. The goal is refined into several questions that usually break down the issue into its major components. Questions try to characterize the object of measurement (product, process, or resource) with respect to a selected quality issue, and to determine its quality from the selected viewpoint. Each question is then refined into metrics, either objective or subjective. Objective metrics include the data that depend only on the object that is being measured and not on the viewpoint from which they are taken. Subjective metrics depend on both the object that is being measured and the viewpoint from which they are taken. The same metric can be used to answer different questions under the same goal. Several GQM models can have questions and metrics in common.

The Goal Question Metric approach is a mechanism for defining and interpreting operational and measurable software. It can be used in isolation or within the context of a more general approach to software quality improvement.

The Goal-Question-Indicator-Measure (GQIM) Model

The Goal-Question-Indicator-Measure (GQIM) model [30] was proposed as part of a goal-driven process that draws extensively on ideas of Basili and

Rombach. The emphasis throughout goal-driven measurement is on gathering information that helps people achieve their business goals, and on maintaining traceability from measures back to business goals.

The goal-driven measurement process is based on 3 percepts as described below:

- Measurement goals are derived from business goals.
- Evolving mental models provide context and focus.
- GQIM translates informal goals into executable measurement structures.

The GQIM process model has 10 steps as listed below. The process begins with identifying business goals and breaking them down into manageable sub-goals. It ends with a plan for implementing well-defined measures and indicators that support the goals.

1. Identify your business goals.
2. Identify what you want to know or learn.
3. Identify your sub-goals.
4. Identify the entities and attributes related to your sub-goals.
5. Formalize your measurement goals.
6. Identify quantifiable questions and the related indicators that you will use to help you achieve your measurement goals.
7. Identify the data elements that you will collect to construct the indicators that help answer your questions.
8. Define the measures to be used, and make these definitions operational.
9. Identify the actions that you will take to implement the measures.
10. Prepare a plan for implementing the measures.

The goal can be initiated at any organizational level, and the output of Step-1 is a sorted checklist of business goals (i.e., management goals, development goals, and maintenance goals, etc) along with their definitions. At Step-2, it is identified what is needed to know in order to understand, assess, predict, or improve the activities related to achieving goals by asking questions. Grouping related questions helps identifying sub-goals at Step-3. The questions about the entities (inputs, outputs, activities, or internal artifacts) are identified and grouped to specify the issues they address, and the groupings of issues and questions translate naturally into candidate sub-goals. At Step-4, each question from Step-2 is examined, entities implicit in the question are identified, and appropriate attributes associated with each entity are listed. The attributes are the candidates for the things that should be measured. Measurement goals are formalized at Step-5, including the descriptions for object of interest, purpose, perspective, and environment. Measurement goals should be traced back to the subgoals and business goals to show that they are consistent with the business objective. At Step-6, quantifiable questions related to each measurement goal are identified, and sketches for displays (indicators) that will help to address identified questions are prepared. An indicator is a display of one or more measurement results that is designed to communicate or explain the significance of those results to the user. Indicators are useful because seeing how measurement data will be displayed helps clarify exactly what must be measured. Data elements that must be collected to construct the indicators are identified at Step-7. Identifying data elements involves preparing a list of data items (attributes) as well as preparing a checklist cross-referencing data items and indicators (i.e., which data element is used by which indicator). At Step-8, measures to be used are defined clearly. A measure definition is a semi-formal specification for the object to be measured, and is extremely useful to clarify the implicit assumptions, what is included and what is not in the measurement. Step-9 is to assemble information about the current status and use of the measures, so as to prepare a plan for implementing defined measures through analysis (fact finding), diagnosis (evaluation), and action (solution finding). Analysis means identifying the measures that the organization is using now and understanding how it is collecting them. Diagnosis means evaluating the data elements that the organization is collecting now, determining how well they meet the needs of new measurements, and proposing appropriate actions. Action means translating the results of the analysis and diagnosis into implementable steps. At Step-10 a measurement implementation plan is prepared based on analysis, diagnosis and actions.

The GQIM model describes an adaptable process that teams and individuals can use to identify and define measures that provide insight into their own management issues. Intended audiences include program managers, project managers, process managers, process improvement teams, and measurement teams.

The Model Measure Manage Paradigm (M^{3}P)

The Model Measure Manage Paradigm (M^{3}P) [31], as an extension of the Quality Improvement Paradigm (QIP) and Goal-Question-Metric (GQM) approach, addresses the lack of well-defined links between the numerical data and the surrounding development business context (Figure 2.3).

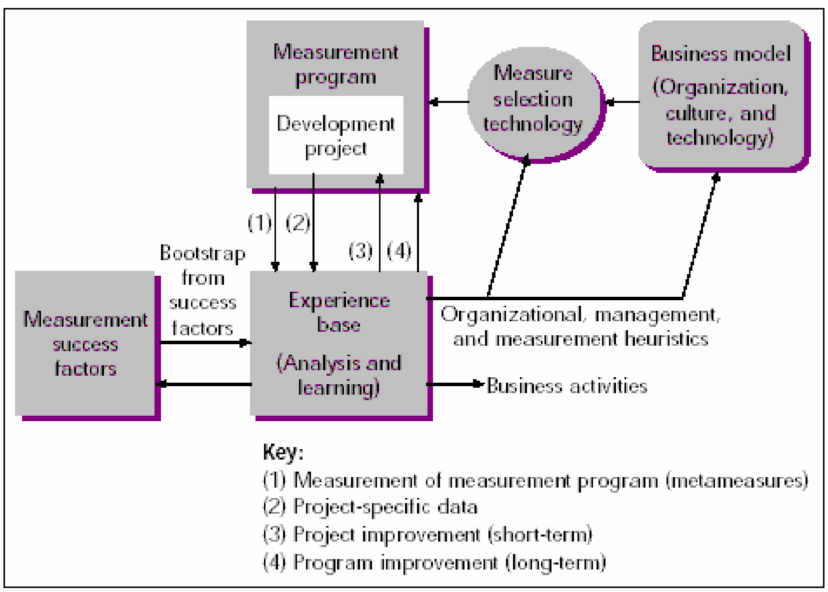


Figure 2.3 The Model Measure Manage Paradigm

QIP, usually coupled with GQM, is stated as a useful tool for decomposing goals into specific measurement requirements; however, it is criticized for often leaving important environmental and measurement issues implicit rather than explicit (e.g., how the top-level goals relate to business imperatives). M^{3}P extends the QIP framework by providing additional features designed to reflect known measurement program success factors, and to support data measurement, analysis, and interpretation. M^3P incorporates GQM as an explicit measure selection technology, which is a means for selecting measures that readily illuminate and support the achievement of business and development goals. M^3P can readily feed measurement data back into the empirical model as shown in the figure.

The M^3P framework helps companies identify the necessary prerequisites to measurement to maximize the relationships between the empirical model and the numerical model. In order to exploit the M^3P framework, developers must progress from prerequisites (business imperatives) to measurement program design, implementation, utilization, and review in an orderly fashion, as described below:

1. Understand the business strategy.
2. Identify business goals, sub-strategies, risks, and tactics that depend on successful software development, use, and support.
3. Determine the critical success factors.
4. Define specific software development goals, based on the first three steps.
5. Pose questions.
6. Identify and define measures.
7. Set up the program: Generate detailed procedures and define reports (for all stakeholders).
8. Regularly review the program by revisiting the above seven steps.

The M^3P mediated progression from high-level business goals to measurement program is stated to be best handled via a series of facilitated workshops. As the measurement program details emerge, they must be checked for viability. This is achieved by the role of “measurement success factors” box in Figure 2.3, based on the Jeffery/Berry success-factor framework. It is reported from the experiences that values derived using this framework correlated well with the success or failure of the organizations studied, and that the framework could serve as a measurement planning and implementation checklist and thus predict the likelihood of success of the emerging measurement program.

The M^3P approach applies to both large and small individual projects as readily as it does to whole organizations. The detailed management and timing of program planning, implementation, and review largely depend on a company's individual culture and management style.

Practical Software Measurement (PSM)

Experience across a wide range of software development and maintenance projects suggests two key characteristics of a successful measurement program: The collection, analysis, and reporting of measurement data that relates directly to the information needs of the project decision makers; and, a structured and repeatable measurement process that defines project measurement activities and related information interfaces. Practical Software Measurement (PSM), based on years of experience of dozens of organizations, has been proposed to address these two key characteristics [41]. PSM addresses the development of a project measurement information structure using the Measurement Information Model, and describes measurement activities and tasks using the Measurement Process Model.

The Measurement Information Model is a mechanism for linking defined information needs to the project's software processes and products, that is, to the entities that can actually be measured. It helps to define the information needs of the project decision makers, and focuses measurement-planning activities on the selection and specification of the most appropriate software measures to address those needs. As the measures are implemented and data are collected, the Measurement Information Model structures the measurement data and associated analysis into structured information products. These information products integrate the measurement results with established decision criteria, and present recommendations to project decision makers on alternative courses of action. PSM makes use of seven common information categories to facilitate the identification and prioritization of a project's specific information needs, which are schedule and progress, resources and cost, product size and stability, product quality, process performance, technology effectiveness, and customer satisfaction.

The Measurement Process Model works in conjunction with the Measurement Information Model, to provide an application framework for implementing measurement on a project (Figure 2.4). The model is built around a typical “Plan-Do-Check-Act” management sequence, adapted to support measurement-specific activities and tasks. The Measurement Process Model includes four primary activities, each of which is essential to successful measurement implementation:

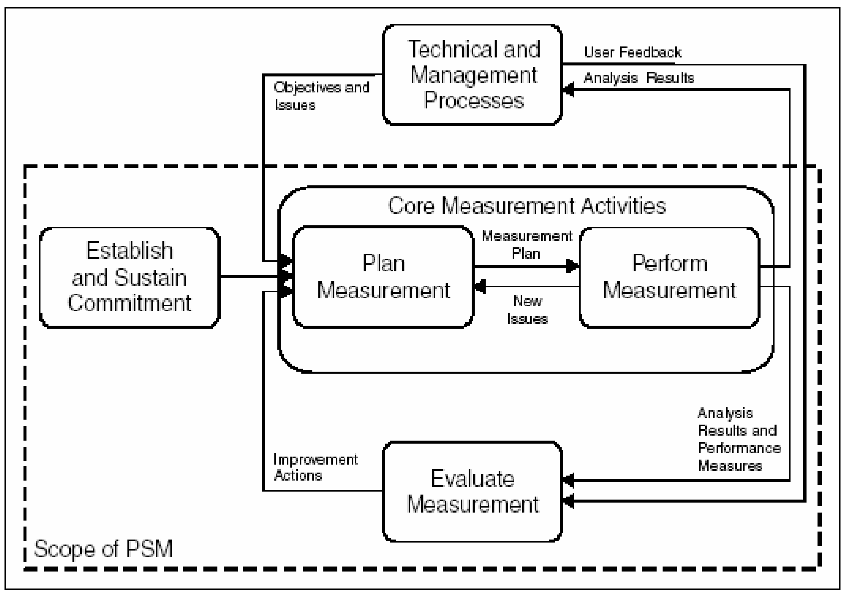


Figure 2.4 The Measurement Process Model of PSM

- Plan Measurement activity encompasses the identification of project information needs and the selection of appropriate measures to address these needs using the Measurement Information Model. Its output is a well-defined measurement approach that directly supports the project’s information needs.
- Perform Measurement activity encompasses the collecting and processing of measurement data. It implements the measurement plan and produces the information products necessary for effective measurement-based decision-making.

- Evaluate Measurement activity applies measurement and analysis techniques to the measurement process itself. It ensures that the project measurement approach is continually updated to address current information needs and promotes an increasing maturity of the project and organizational measurement process.
- Establish and Sustain Commitment activity ensures that measurement is supported both at the project and organizational levels. It provides the resources and organizational infrastructure required to implement a viable measurement program.

The Measurement Process Model is iterative by design and it is defined to be tailored to the characteristics and context of a particular project and to be adaptable to changing project information and decision requirements.

2.4. Literature on Measurement Practices

ISO/IEC TR 15504: IT – Software Process Assessment

ISO/IEC TR 15504 [3] provides a framework for the assessment of software processes, and includes a process dimension mapped against a capability dimension. The assessment characterizes the current practice within an organizational unit in terms of the capability of the selected processes. The process dimension includes life cycle processes under basic categories as customer-supplier, engineering, support, management, and organization. The capability dimension includes 6 levels (from 0 to 5), each composed of one or more process attributes.

ISO/IEC TR 15504 defines indicators of process capability for management practices which are defined under process attributes in the capability dimension. The indicators are given in Table 2.2.

Since the process control comes with Level 4 (specifically with process attribute 4.2), indicators of management practices up to that level give idea about the characteristics of a process which is subject to statistical control.

Table 2.2 Capability Levels and Process Attributes in ISO/IEC TR 15504

Capability Level	**Process Attribute**	**Management Practice (Indicator Class)**
1: Performed	1.1: Process performance	1.1.1: Identify input and output work products. 1.1.2: Ensure that the scope of work is identified for process execution and for the work products to be used and produced by the process. 1.1.3: Ensure that base practices are implemented, producing work products which support achievement of the defined process outcomes.
2: Managed	2.1: Performance management	2.1.1: Identify the objectives for the performance of the process (for example; time scale, cycle time, and resource usage). 2.1.2: Plan the performance of the process according to the identified objectives by identifying activities of the process, the expected time schedule, and allocation of resources for each activity. 2.1.3: Plan and assign the responsibility and authority for developing the work products of the process. 2.1.4: Manage the execution of the activities by continued tracking and re-planning to produce work products that meet the defined objectives.
	2.2: Work product management	2.2.1: Identify the requirements for the work products, including both functional and non-functional aspects. 2.2.2: Manage the documentation, configuration management, and change control of the work products. 2.2.3: Identify and define any work product dependencies. 2.2.4: Manage the quality of work products to ensure that they meet their functional and non-functional requirements.
3: Established	3.1: Process definition	3.1.1: Identify the standard process that supports the execution of the managed process and provides documented guidance on tailoring. 3.1.2: Implement and/or tailor the standard process to obtain a defined process appropriate to the process context. 3.1.3: Gather process performance data so that the behavior of the defined process can be understood. 3.1.4: Establish and refine the understanding of the process behavior by using process performance data. 3.1.5: Refine the standard process.
	3.2: Process resource	3.2.1: Identify and document the roles, responsibilities, and competencies required to support the implementation of the defined process. 3.2.2: Identify and document the process infrastructure requirements to support the implementation of the defined process. 3.2.3: Provide, allocate, and use the resources to support the performance of the defined process. 3.2.4: Provide, allocate, and use an adequate process infrastructure to support the performance of the defined process.

Table 2.2 Capability Levels and Process Attributes in ISO/IEC TR 15504 (cont'd)

4: Predictable	4.1: Measurement	4.1.1: Identify product and process goals and measures which support the achievement of the relevant business goals. 4.1.2: Collect the specified product and process measures through performing the defined process. 4.1.3: Analyze trends in the performance of the process across the organization. 4.1.4: Measure the process capability and maintain it within the defined limits across the organization.
	4.2: Process control	4.2.1: Identify suitable measurement techniques, appropriate to the process context, to support process and product improvement. 4.2.2: Collect measures and identify process control parameters in order to perform analysis. 4.2.3: Control the process performance using the analysis measures to identify actions to maintain control and/or implement improvement.
5: Optimizing	5.1: Process change	5.1.1: Identify changes to the standard process definition on the basis of a quantitative understanding of the process. 5.1.2: Assess the impact of all proposed changes against the defined product and process goals of the defined and standard processes. 5.1.3: Define an implementation strategy for the approved change, ensuring that any disruption to the process performance is understood and acted upon. 5.1.4: Implement the approved changes to the affected process according to the implementation strategy. 5.1.5: Evaluate the effectiveness of process change on the basis of actual performance against the defined product, process, and business goals, making adjustments as needed.
	5.2: Continuous improvement	5.2.1: Define the process improvement goals for the process that support the relevant business goals of the organization. 5.2.2: Analyze the source of real and potential problems in the current process, identifying improvement opportunities in a systematic and proactive manner to continuously improve the process. 5.2.3: Implement changes to selected areas of the tailored process according to the implementation strategy. 5.2.4: Validate the effectiveness of process change on the basis of actual performance against process and business goals and feedback to the standard process definition.

The Capability Maturity Model Integrated

CMMI defines Measurement and Analysis process area at level 2, which requires software projects define their specific information needs and metrics that will serve for these needs, as well as metric data collection, analysis, and sharing procedures. CMMI also defines Quantitative Project Management process area at level 4, which proposes quantitative management of the project's defined process.

These two process areas show that although CMMI encourages measurement practices at level 2, it still postpones process control until level 4.

(Level-2) Measurement and Analysis:

The purpose of this process area is to develop and sustain a measurement capability that is used to support management information needs. Measurement capability may be integrated into individual projects or other organizational functions (e.g., quality assurance). The initial focus for measurement activities is at the project level. However, a measurement capability may prove useful for addressing organization- and/or enterprise-wide information needs.

Generic and specific goals (GG and SG) as well as generic and specific practices of the process area are given in Table 2.3. They are the specific goals and practices that define the tasks specific to measurement and analysis.

Table 2.3 CMMI Measurement and Analysis Process Area – Goals and Practices

Goal	**Practice**
SG 1 Align Measurement and Analysis Activities	SP 1.1 Establish Measurement Objectives
	SP 1.2 Specify Measures
	SP 1.3 Specify Data Collection and Storage Procedures
	SP 1.4 Specify Analysis Procedures
SG 2 Provide Measurement Results	SP 2.1 Collect Measurement Data
	SP 2.2 Analyze Measurement Data
	SP 2.3 Store Data and Results
	SP 2.4 Communicate Results
GG 2 Institutionalize a Managed Process	GP 2.1 (CO 1) Establish an Organizational Policy
	GP 2.2 (AB 1) Plan the Process
	GP 2.3 (AB 2) Provide Resources
	GP 2.4 (AB 3) Assign Responsibility
	GP 2.5 (AB 4) Train People
	GP 2.6 (DI 1) Manage Configurations
	GP 2.7 (DI 2) Identify and Involve Relevant Stakeholders
	GP 2.8 (DI 3) Monitor and Control the Process
	GP 2.9 (VE 1) Objectively Evaluate Adherence
	GP 2.10 (VE 2) Review Status with Higher Level Management
GG 3 Institutionalize a Defined Process	GP 3.1 Establish a Defined Process
	GP 3.2 Collect Improvement Information

The Measurement and Analysis process area involves the following:

- Specifying the objectives of measurement and analysis such that they are aligned with identified information needs and objectives;

- Specifying the measures, data collection and storage mechanisms, analysis techniques, and reporting and feedback mechanisms;
- Implementing the collection, storage, analysis, and reporting of the data;
- Providing objective results that can be used in making informed decisions and taking appropriate corrective actions.

(Level-4) Quantitative Project Management:

The purpose of this process area is to quantitatively manage the project's defined process to achieve the project's established quality and process-performance objectives. Generic and specific goals (GG and SG) as well as generic and specific practices of the process area are given in Table 2.4.

Table 2.4 CMMI Quantitative Project Management Process Area – Goals and Practices

Goal	**Practice**
SG 1 Quantitatively Manage the Project	SP 1.1 Establish the Project's Objectives
	SP 1.2 Compose the Defined Process
	SP 1.3 Select the Sub-processes that Will Be Statistically Managed
	SP 1.4 Manage Project Performance
SG 2 Statistically Manage Sub-process Performance	SP 2.1 Select Measures and Analytic Techniques
	SP 2.2 Apply Statistical Methods to Understand Variation
	SP 2.3 Monitor Performance of the Selected Sub-processes
	SP 2.4 Record Statistical Management Data
GG 3 Institutionalize a Defined Process	GP 2.1 (CO 1) Establish an Organizational Policy
	GP 3.1 (AB1) Establish a Defined Process
	GP 2.2 (AB 2) Plan the Process
	GP 2.3 (AB 3) Provide Resources
	GP 2.4 (AB 4) Assign Responsibility
	GP 2.5 (AB 5) Train People
	GP 2.6 (DI 1) Manage Configurations
	GP 2.7 (DI 2) Identify and Involve Relevant Stakeholders
	GP 2.8 (DI 3) Monitor and Control the Process
	GP 3.2 (DI 4) Collect Improvement Information
	GP 2.9 (VE 1) Objectively Evaluate Adherence
	GP 2.10 (VE 2) Review Status with Higher Level Management

The Quantitative Project Management process area involves the following:

- Establishing and maintaining the project's quality and process performance objectives;

- Identifying suitable sub-processes that compose the project's defined process based on historical stability and capability data found in process performance baselines or models;
- Selecting the sub-processes of the project's defined process to be statistically managed;
- Monitoring the project to determine whether the project's objectives for quality and process performance are being satisfied, and identifying appropriate corrective action;
- Selecting the measures and analytic techniques to be used in statistically managing the selected sub-processes;
- Establishing and maintaining an understanding of the variation of the selected sub-processes using the selected measures and analytic techniques;
- Monitoring the performance of the selected sub-processes to determine whether they are capable of satisfying their quality and process-performance objectives, and identifying corrective action;
- Recording statistical and quality management data in the organization's measurement repository.

The quality and process-performance objectives, measures, and baselines identified above are developed as described in the Organizational Process Performance process area. Subsequently, the results of performing the processes associated with the Quantitative Project Management process area (e.g., measurement definitions and measurement data) become part of the organizational process assets referred to in the Organizational Process Performance process area.

To effectively address the specific practices in this process area, the organization should have already established a set of standard processes and related organizational process assets, such as the organization's measurement repository and the organization's process asset library, for use by each project in establishing its defined process. The project's defined process is a set of sub-processes that form an integrated and coherent life cycle for the project. It is established, in part, through selecting and tailoring processes from the organization's set of standard processes.

Sub-processes are defined components of a larger defined process. For example, a typical organization's development process may be defined in terms of sub-processes such as requirements development, design, build, test, and peer review. The sub-processes themselves may be further decomposed as necessary into other sub-processes and process elements.

One essential element of quantitative management is having confidence in estimates (i.e., being able to predict the extent to which the project can fulfill its quality and process-performance objectives). The sub-processes that will be statistically managed are chosen based on identified needs for predictable performance.

Another essential element of quantitative management is to understand the nature and extent of the variation experienced in process performance, and recognizing when the project's actual performance may not be adequate to achieve the project's quality and process performance objectives.

Statistical management involves statistical thinking and the correct use of a variety of statistical techniques, such as run charts, control charts, confidence intervals, prediction intervals, and tests of hypotheses. Quantitative management uses data from statistical management to help the project predict whether it will be able to achieve its quality and process-performance objectives and identify what corrective action should be taken.

This process area applies to managing a project, but the concepts found here also apply to managing other groups and functions. Applying these concepts to managing other groups and functions may not necessarily contribute to achieving the organization's business objectives, but may help these groups and functions control their own processes.

ISO/IEC 15939 - Software Measurement Process

This international standard [42] contains a set of activities and tasks that comprise a software measurement process that meets the specific needs of software organizations and projects. It defines the activities and tasks necessary to successfully identify, define, select, apply, and improve software measurement within an overall project or organizational measurement structure (Figure 2.5). It is intended to use by software suppliers and acquirers. The measurement process should be appropriately integrated with the organizational quality system.

The standard does not assume or prescribe an organizational model for measurement. The users should decide whether a separate measurement function is necessary within the organization, whether the measurement function should be integrated within individual software projects or across projects, and etc. based on the current organizational structure, culture, and prevailing constraints. In addition, it is not intended to prescribe the name, format, or explicit content of the documentation to be produced, and leaves these decisions to its users.

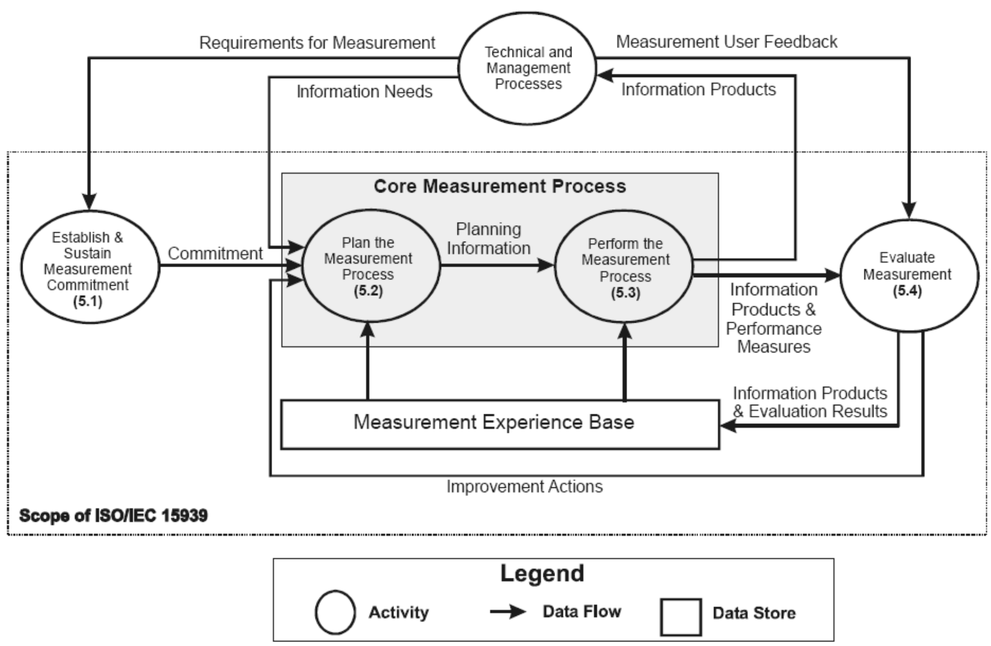


Figure 2.5 ISO/IEC 15939 Software Measurement Process

The 2001 High Maturity Workshop

In March of 2001, the Software Engineering Institute (SEI) hosted a workshop for high maturity organizations to better understand practices that characterize

CMM for Software Level 4 and 5 organizations. Topics of discussion included practices described in the Software CMM as well as other practices that have a significant impact in mature organizations. Important themes included statistical process control for software, the reliability of Level 4 and 5 assessments, and the impact of the CMMI effort. Additional topics solicited from the participants included measurement, Six Sigma, Internet speed and process agility, and people and cultural issues. This report contains overviews of more than 30 high maturity organizations and the various working group reports from the workshop.

Below are the findings from workgroup on measurement (WG 1.1):

Summary of observations and hypothesis:

- Do not try to standardize measures across too large an organization. A better approach might be to identify common categories with sepcific common measures.
- Changing definitions of measures degrades the utility of historical databases, baselines, and parametric models; and, makes automation of data collection difficult. (This finding underlies the importance of making metric definition right the first time.) E.g., measurement definitions may influence how programmers choose to format their code.
- Across the organizations represented in the working group, there was a great deal of commonality in the data collected: cost, schedule, effort, size, changes, defects, etc.
- Managing the culture change as an organization seeks to move to CMM Level-4 seemed to be widespread issue. CMM literacy is somehow dangerous, since it gives lower maturity projects a reason not to focus on measurement. E.g., I just want to be Level-3, so I do not have to measure...
- CMM Level-3 was inherently unstable. Advancing organizations apply the data or derive additional measures to control and improve processes. Groups that do not take this next step tend to stop doing other than basic cost/schedule measuring and hence regress. “Just enough” CMM literacy to be dangerous – measurement is not just for high-maturity organizations. Measurement is

essential to knowing where you are with respect to your program and process goals, and this understanding should begin with the first key process area at the lower maturity levels.

Summary of pitfalls and false starts:

- Choosing the wrong measures is one of the biggest potential pitfalls. Follow "why → what → how" sequence while defining metrics. Be aware that what worked for another organization may not be directly applicable to you.
- Disconnects between project goals and process improvement goals can cause perpetual confliction of priorities, and prohibit understanding and progress. E.g., involve stakeholders; beware the SEPG (Software Engineering Process Group) trying to do it all themselves and not getting involved in projects (as the maturity increases, SEPG becomes facilitator).

Summary of recommendations:

- Do it "with" them, not "to" them.
- Start with small, focused efforts.
- Integrate project measures and business objectives.
- Re-visit the basis, and review purpose and need for each measure.
- Automate collection and analysis as possible.
- Address change management, including people issues.

Below are the findings of workgroup on statistical techniques (WG 2.1):

Summary of observations and hypothesis:

- Managers generally do not have an understanding of statistical methods.
- Metrics and process improvement activities must be tied to business results (application of the GQM approach required).
- Most examples for using statistical techniques involve inspection/defect data. How other types of data (e.g., cost, schedule, reliability) can be analyzed to improve reliability?

- CMM Level-4, Level-5, and Six Sigma need to be integrated. Six Sigma can be a new approach.
- CMM Level-3 measures are often not adequate to support Level-4. Level-3 measures generally stay at the project phase level, and do not provide the granularity needed at Level-4. The focus for these measures needs to go beyond just cost and schedule, and quality should be measured at Level-2 and Level-3. Measures that provide valuable insight and control should be used by all organizations regardless of maturity.
- SPC should be treated as "one tool" in a process improvement toolkit. Generally, classic SPC should be used in situations where sources of variance are better controlled. E.g., control charts done badly are worse than not doing SPC at all, since it can give a false impression of process stability (large variation with little predictive value).
- Statistical techniques proven to be useful outside of software development, but rarely used by CMM Level-4 and Level-5 companies include: multivariate methods for variance, non-parametric statistics for unusual distributions, reliability and statistically based testing for determining operational performance profiles, and Bayesian methods.
- Process performance baselines need to be maintained at project as well as at organization levels. Maintain process capability baselines for: productivity, delivered defects, in-process defects (defect profiles by phase), and defects/LOC and defects/hr for each type of peer review.

Summary of recommendations for high maturity organizations:

- Let the data and objectives determine the statistical methods used.
- Set quantitative objectives tied to business goals.
- Simplify the presentation of statistical results.
- Use data to gain understanding and control, and to guide improvement.
- Learn about Six Sigma and the tools it offers for CMM Level-4 and Level-5.

2.5. Relation of the Literature to Our Study

When we look at the literature on SPC for software, we see that there is a gap between the implementations (explained in section 2.2.1) and the guidelines (described in section 2.2.2). The implementations are mostly specific to their own cases, and therefore do not explain their practices as possible to follow by similar organizations. Guidelines, on the other hand, provide a list of things that should be done or not be done, and underline key points to succeed. There is a lack of defined methods that will tie generic guidelines to specific implementations, and the need for this lack directed us to define a model that will close the gap.

Both literature on measurement approaches and models (provided in section 2.3) and literature on measurement practices (provided in section 2.4) contributed to the development of the model. Literature on measurement approaches and models showed direction primarily for goal-oriented measurement, measurement theory, and metric definition. Literature on measurement practices enabled us to identify the relationship between measurement and process maturity and to determine the practices to search in a typical measurement implementation. Although being general, guidelines on SPC for software (described in section 2.2.2) provided necessary background on concepts like process performance, stability, and capability. These concepts enabled us to understand internal and external process factors that contribute to statistical control.

Since we wanted the model to be practical and usable for any organization, we decided it to be an assessment model to test the applicability of SPC for a software process and its metrics. We aimed that the assets of the model primarily support understanding of process components and the context in which the data is generated. This understanding is crucial for statistical analysis and interpretation of analysis results. We also aimed that the assets show guidance to select process metrics to use for statistical analysis.

As a result, our model would rely upon “known” concepts of process and measurement (like rational sampling, measurement scales, measurement practices, etc.), and would propose a “new” assessment method to specifically guide any company to implement SPC for its own system/software development process.

CHAPTER 3

AN ASSESSMENT MODEL FOR STATISTICAL PROCESS CONTROL

We developed an assessment model to evaluate the applicability of SPC for software process. The basic intention behind the model was to identify guidelines to direct SPC implementation. Accordingly, we considered two basic requirements for SPC implementation, and focused on finding ways to resolve the difficulties brought by these requirements for an emergent organization while developing the model: Rational sampling of process executions and data, and metric data utilization (or suitability) for statistical analysis.

3.1. Model Components

The first requirement is the rational sampling of process executions and data. The purpose of rational sampling is to obtain and use data that are representative of the performance of the process with respect to the issues being studied. If we can consider that observations are made under essentially the same conditions and that differences between the measurements are primarily due to common cause variation, then we are very likely that we rationally group the observations [35].

Since we want to sample process executions as being from a single and constant system of chance causes, we developed a clustering method based on process attributes such as inputs, outputs, activities, roles, and tools and techniques. The relation of these attributes to the process is given in Figure 3.1. If repetitions of a process show similarity in terms of these attributes, then we assume that the process is consistently performed among its executions. Process attributes are briefly described below:

- Input: An entity that have been entered into the process or expended in its operation to achieve one or more outputs. The process has a number of inputs to each execution.
- Output: An entity that have been produced by the process or created in its operation to satisfy process purpose. The process has a number of outputs from each execution.
- Activity: A distinct step within the process, when completed, supports transformation of input(s) into output(s) to achieve process purpose. The process has a number of activities that are carried out within each execution.
- Role: The actions assigned to or required of a person or group to carry out the activities within the process. The process allocates responsibility to a number of roles that participates in one or more process activities.
- Tools and Techniques: An implement used in or a practical method applied to some particular activity to support its completion. The process holds a number of tools and techniques that are used in one or more process activities.

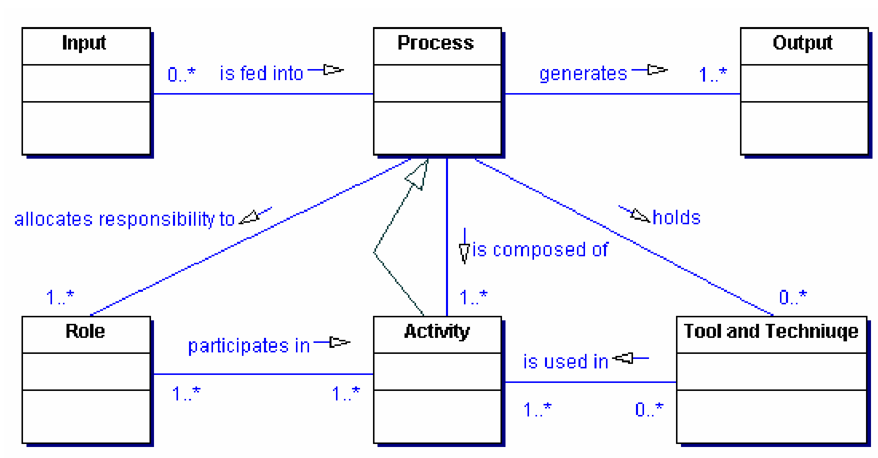


Figure 3.1 Process Attributes used for Rational Sampling

Process consistency is assessed for similarity in process attribute values of process executions. We record the attribute values of each execution on a form, and to compare the similarity of these recorded values on a matrix. Ideally it is desirable that the process has a unique version in execution. When it has, it is more likely that we have a single and constant system of chance causes. However, a process might have several versions (and therefore mixed systems for chance causes) in execution as well. The idea behind process consistency assessment as basis for rational sampling is to identify, if any, these differing versions of a process in execution. We should note that the assessment of consistency can also be performed for a specific activity, since the activity "is-a" process as shown in Figure 3.1. This means that if the analysis we perform at process level does not provide expected insight for process understanding and control, we then can apply the approach at a lower (activity) level.

The second requirement is metric utilization. This includes elaboration of basic measurement practices as well as metric data existence and characteristics. Measurement practices should be performed for a specific purpose [28][29][30] and, metrics should be uniquely understood to enable consistent implementation. Unique understanding (mostly enabled by constructing operational definitions) requires three criteria: communication, repeatability, and traceability [35]. The traceability requirement is especially important to assessing and improving process performance. Because measures of performance can signal process instabilities, it is important that the context and circumstances of the measurement be recorded. This helps identifying assignable causes of the instabilities. There are studies that define procedures for successfully implementing measurement practices and for incorporating measurement capability into the projects of an organization [2][30][41][42]. The CMMI [2] for example, introduces Measurement and Analysis process area at maturity level 2, and recommends practices for defining data collection, storage, analysis, and reporting. Existence and implementation of these practices can be questioned for a specific project or organization to determine the utilization of existing metrics and data. Also, there are high-maturity companies that developed the factors to consider for measurement evaluation and to determine what measures to select for their specific use [43].

To evaluate metric utilization, we identified a number of metric usability attributes, and developed questionnaires based on these attributes for base and derived metrics separately. Table 3.1 lists and explains these attributes. Questionnaires include a rating system based on the answers of questions, and accordingly, evaluate the usability of a specific metric for applying SPC. A metric must satisfy the scale type requirement (absolute or ratio) [1] and have enough data points to use (20 at a minimum) [38] as specified by the first two attributes. Verifiability and dependability of metric data significantly contribute to the confidence in data analysis results. Data verifiability is related with the consistency in metric data recording and storage among executions. Data dependability requires all metric data be recorded as close to its source with accuracy and precision. The awareness of data collectors on metric data (why it is collected, how it is utilized, etc.) plays a significant role in data dependability. The last two attributes, data normalizability and data integrability, are related with the usefulness of a metric and should be satisfied if we expect SPC analysis provide more insight for process understanding and improvement.

Table 3.1 Metric Usability Attributes used for Evaluating Metric Utilization

Metric Usability Attribute	**Explanation**
Metric Identity	Metric should be identified including entity and attribute to measure; scale type, unit, formula; and data type and range. Included in the identity is the scale type of the metric. Nominal and ordinal scale metrics cannot be used for control charting.
Data Existence	For any analysis, there should be measurement data. For control limits to be calculated reliably there should be at least 20 data points.
Data Verifiability	Metric data should be recorded at the same place in the process, by the same responsible body, and using the same method every time.
Data Dependability	Metric data should be recorded and stored as it is generated to ensure accuracy and precision; and be collected for a specific purpose. Feedback mechanisms should exist and be known by data collectors regarding data analysis and reporting.
Data Normalizability	Metric data can be normalized with a parameter or with another metric. Normalizing metric-A with a parameter-P provides comparable values of metric-A in terms of the parameter-P. Normalized metrics provide more insight in terms of statistical analysis (e.g., normalizing number of defects in a product with product size).
Data Integrability	Metric data can be integrated at project or organization levels. In practice, metric data should be integrated from individual level up to organization level for the results of statistical analysis to be effective organization-wide.

3.2. Assessment Process

The assessment process to follow when applying the model is given in Figure 3.2.

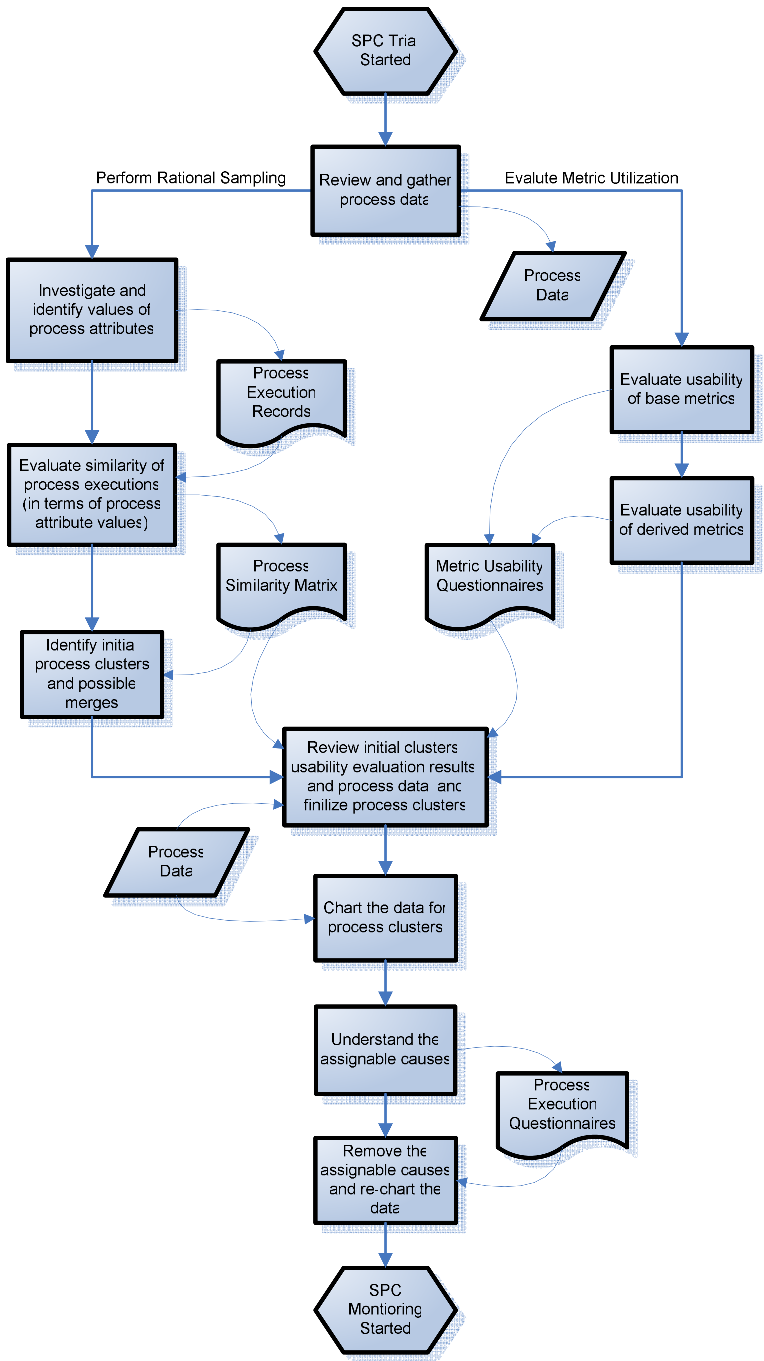


Figure 3.2 The Assessment Process

The first step of the assessment process is reviewing and gathering process data typically in a data file. Data should be consolidated in time sequence and in a form that is appropriate for comparison among different projects and product types. During consolidation, traceability should be established between process executions and data, typically by giving the same identifier to both. The data of process executions having missing, incomplete, or invalid data points should be excluded.

The flow at the left side of the figure is for performing rational sampling. We investigate and identify the values of process attributes for process executions by filling out process execution records. If the study is retrospective then we sample several executions from past process performances and fill a record for each. A merged list of values is built from process attribute values of sampled executions on records and entered into process similarity matrix for verification against entire set of process executions. The list on the matrix is extended during verification when a new value shows up.

If the study is prospective, a process execution record is filled when a new instance of the process is being executed. This increases our confidence on the values of process attributes for a process execution. Another difference in a prospective study is that we fill a process execution questionnaire for each instance of the process in execution and at the same time we fill a process execution record (not while searching for the assignable causes later in the process as shown in Figure 3.2). This is because we want to capture the external factors affecting the process execution more timely, and have the chance of identifying likely assignable causes in advance.

The last step of the flow at the left side of Figure 3.2 as basis for rational sampling is identifying initial process clusters and possible merges among them by analyzing the process similarity matrix. How we analyze the matrix and identify the clusters is described in part 3.3, and will not be repeated here.

The flow at the right side of the figure is for evaluating metric utilization. First, usability of each base metric and then usability of each derived metric is evaluated by filling a metric usability questionnaire, and calculating regarding metric usability result. How we derive the metric usability result is described in part 3.3, and will not be repeated here.

After we identify initial process clusters and evaluate usability of process metrics, we use the knowledge we gathered so far as well as process data to finalize process clusters and metrics as basis for control charting. This is where the flows at left and right sides join in Figure 3.2. Here we review initial process clusters and possible merges among them, the number of data points for each process cluster, and the usability status of process metrics; and identify the resulting *process cluster-process metric* pairs to chart. During review, we can decide to exclude process clusters with few data from the study or to merge the most similar ones to increase the number of data points. Our model recommends charting the data for process metrics that are evaluated as “usable” for statistical analysis; however, it might be a good idea to chart the data for the metrics that are evaluated as “not usable” to validate (or invalidate) the model’s recommendation. It is better to review the number of data points per process metric basis since there may be missing data points. We suggest composing data sets including the data from all process clusters for each process metric, and then eliminating missing data points in each data set.

We then separately put the data for *process cluster-process metric* pairs on control charts, and watch for the out-of-control points. In a retrospective study, we fill process execution questionnaire for each out-of-control point to understand the assignable causes if any. In a prospective study, we review previously filled process execution questionnaires to understand the assignable causes. Additionally, we suggest performing interviews with process performers to detect any reasons for out-of control points, or potential assignable causes that the process execution questionnaires cannot catch. After removing data points regarding the assignable causes at each chart, we re-chart the data for each *process cluster-process metric* pair and watch if the data on the chart is under control. Here is the place to judge whether our approach helped us in starting SPC. If a chart regarding a process cluster-process metric pair validates the findings of the assessment model, then SPC monitoring begins for that pair.

3.3. Assessment Assets

We developed several assets for use in the assessment to perform rational sampling and to evaluate metric utilization. *Process execution record* together with *process similarity matrix* is utilized to identify process clusters as basis for rational sampling. *Metric usability questionnaires* are used to evaluate metrics' usability for SPC, and *process execution questionnaire* is used to investigate assignable causes for an out-of-control point on a control chart. The following paragraphs describe these assets. Original copies of these assets are provided in Appendix-A.

Process Execution Record is a form used to capture the instant values of process attributes for a process execution. Actual values of inputs, outputs, activities, roles, and tools and techniques for a specific process execution are recorded on the form (Figure 3.3). Recorded values are used to identify the merged list of process attribute values which are entered into Process Similarity Matrix for verification.

Process Name:		Recorded On:	
Process Execution No:		Recorded By:	

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1		

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1		
2		

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1		
2		

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1		
2		

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
1		
2		

Figure 3.3 Process Execution Record

Process Similarity Matrix is a spreadsheet used to verify process attribute values against process executions. Process attribute values are recorded into the rows of the matrix vertically and process execution numbers are recorded into the columns of the matrix horizontally. By going over process executions, the values of process attributes are questioned and marked if applicable for each process execution (Figure 3.4). The completed matrix helps us to see the differences among process executions in terms of process attribute values, and enables us to identify rational samples of the process executions accordingly.

			Process Executions																				
Process Attributes			PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	.....
1	**Inputs**																						
	1.1	<Input-1>	o	o	.....																		
	1.2	<Input-2>	o	o	.....																		
2	**Outputs**																						
	2.1	<Output-1>	o	o	.....																		
	2.2	<Output-2>		o	.....																		
3	**Activities**																						
	3.1	<Activity-1>	o	o	.....																		
	3.2	<Activity-2>	o	o	.....																		
	3.3	<Activity-3>	o	o	.....																		
	3.4	<Activity-4>		o	.....																		
4	**Roles**																						
	4.1	<Role-1>	o	o	.....																		
	4.2	<Role-2>	o	o	.....																		
5	**Tools and Techniques**																						
	5.1	<Tools and Techniques-1>	o	o	.....																		
	5.2	<Tools and Techniques-2>		o	.....																		

Figure 3.4 Process Similarity Matrix

We analyze the completed matrix for similarity and differences in process executions. We specifically look for process executions with different attribute values and copy each as a separate cluster as shown in Figure 3.5, while skipping the similar ones. Each column after the column of process attributes in the figure represents a different process cluster. Each process cluster we identify is a rational sample of process executions in terms of process attributes, and ideally we can chart the data for each cluster to see whether the process cluster is under control with respect to a specific metric.

In chapter 2, we stated that a metric should have a purpose if we want to analyze its data and derive some conclusions based on the results. If we expect to take an action based on analysis results, we should know in advance what question we are dealing with or which purpose we want to achieve. This is just the same in control charting. We construct the chart for a specific metric and analyze the results considering the purpose of SPC implementation. Therefore we pay attention to the purpose of applying SPC while identifying possible merges among process clusters (for example, when we lack enough data points) and while investigating the stability of the clusters by our model.

1	**Inputs**							
	1.1	<Input-1>	∘	∘	∘	∘	∘	∘
	1.2	<Input-2>	∘	∘		∘		∘
2	**Outputs**							
	2.1	<Output-1>	∘		∘	∘	∘	
	2.2	<Output-1>	∘	∘	∘	∘	∘	∘
3	**Activities**							
	3.1	<Activity-1>	∘	∘	∘	∘	∘	∘
	3.2	<Activity-2>	∘	∘	∘	∘	∘	∘
	3.3	<Activity-3>	∘	∘	∘		∘	∘
	3.4	<Activity-4>		∘		∘		∘
4	**Roles**							
	4.1	<Role-1>	∘	∘	∘	∘	∘	∘
	4.2	<Role-2>	∘	∘	∘	∘	∘	∘
5	**Tools and Techniques**							
	5.1	<Tools and Techniques-1>					∘	∘
	5.2	<Tools and Techniques-2>				∘		
Process Cluster			A	B	C	D	E	F

A-B	2								
A-C	1	B-C	3						
A-D	2	B-D	2	C-D	3				
A-E	2	B-E	4	C-E	1	D-E	5		
A-F	3	B-F	1	C-F	4	D-F	3	E-F	3

Cluster		Mergable To	Will Be Merged To
B	→	--	--
C	→	A	A
D	→	--	--
E	→	C	A
F	→	B	B

Figure 3.5 Process Clusters and Cluster Distances

To identify possible merges among process clusters, we work on pairs of clusters. We calculate the number of differing attribute values between two clusters, and call this number as “cluster distance”. For example, the distance between the clusters A and B in Figure 3.5 is 2, since the attribute values of these clusters differ for process attributes 2.1 and 3.4. Similarly, cluster distance between process clusters A and C is 1, because the values for attribute 1.2 differ only. We record the distances between the pairs of process clusters in the form of a triangle given in the upper right corner of Figure 3.5. Every row in the *cluster distance triangle* shows us which clusters that a specific process cluster is the most similar to in terms of process attributes. For example, the fifth row of the triangle holds distance values of process cluster F to other clusters. When we have a close look at these values, we see that the distance between the clusters B and F is 1, meaning that B is the most similar cluster for F. When identifying possible merges, we search for the pairs of clusters having a distance of 1. If a row includes the distance values all above 1 (e.g. cluster D in row 3), we concern the related cluster “not mergable” to any other cluster. Therefore by going over the rows of the cluster distance triangle, we identify the clusters with a distance of 1, if any, for each cluster; and record these clusters in a table showing mergable clusters. The table shown below the triangle in the figure provides this information.

We should note that identification of mergable clusters can be done according to the purpose of the metric utilized for statistical control. A metric can measure an attribute of three basic entities: Process, product, and resource [1]. Accordingly, we select a metric either for process management, for product (quality) management, or for resource management. The purpose of the metric that we utilize on a control chart therefore can affect the selection of process attributes used for calculating the cluster distances (e.g., using “inputs” process attribute only for controlling the defectiveness of product types under review in an organization). It can also affect the value of cluster distance allowed for identifying mergable clusters (e.g., not allowing cluster distances while trying to meet customer specification limits set for code defectiveness in a project.)

Metric Usability Questionnaire is a form used to investigate the usability of a process metric in terms of metric usability attributes. The form has two types, for base metrics (Figure 3.6) and derived metrics (Figure 3.7) separately. The form includes a number of questions as indicators of usability attributes. Answers to some questions are informative (shaded under “rating” column of MUQ in the figures) and answers to some are used to rate each usability attribute (expected answers to such questions are given in the rightmost column of MUQ in the figures). A metric usability attribute is rated as a corresponding metric usability factor (MUF) within four ordinal values, based on the answers to its indicators: Fully satisfied (F: %86-100), Largely satisfied (L: %51-85), Partially satisfied (%16-50), and Not satisfied (N: %0-15).

	Please rate each attribute in four scales, based on asnwers to questions as indicators:
Metric Name:	F : Indications of the atribute are fully satisfied (%86-100)
Conceptual Definition:	L : Indications of the atribute are largely satisfied (%51-85)
Assessed On:	P : Indications of the atribute are largely satisfied (%16-50)
Assessed By:	N : Indications of the atribute are not satisfied (%0-15)

Attributes	Indicators	Answers	Rating	Expected Answers
Metric Identity		MUF-1	F	
Q1	Which entity does the metric measure?			
Q2	Which attribute of the entity does the metric measure?			
Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)			Ratio, Absolute
Q4	What is the unit of the metric data?			
Q5	What is the type of the metric data? (integer, real, etc.)			
Q6	What is the range of the metric data?			
Data Existence		MUF-2	F	
Q7	Is metric data existent?			Available > 20
Q8	What is the amount of overall observations?			
Q9	What is the amount of missing data points?			
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)			
Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)			
Data Verifiability		MUF-3	F	
Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)			
Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)			Yes
Q14	Who is responsible for recording metric data?			
Q15	Is all metric data recorded by the responsible body?			Yes
Q16	How is metric data recorded? (on a form, report, tool, etc.)			
Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)			Yes
Q18	Where is metric data stored? (in a file, database, etc.)			
Q19	Is all metric data stored in the same place? (in a file, database, etc.)			Yes
Data Dependability		MUF-4	F	
Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)			
Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)			
Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)			
Q23	Are the frequencies for data generation, recording, and storing different?			No
Q24	Is metric data recorded precisely?			Yes
Q25	Is metric data collected for a specific purpose?			Yes
Q26	Is the purpose of metric data collection known by process performers?			Yes
Q27	Is metric data analyzed and reported?			Yes
Q28	Is metric data analysis results communicated to process performers?			Yes
Q29	Is metric data analysis results communicated to management?			Yes
Q30	Is metric data analysis results used as a basis for decision making?			Yes
Data Normalizability				
Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)			
Data Integrability				
Q32	Is metric data integrable at project level?			
Q33	Is metric data integrable at organization level?			

(a) Metric Usability Questionnaire

Metric Name:		
Conceptual Definition:		
Assessed On:		
Assessed By:		
Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	F	L or F
Metric Usability Result	**F**	L or F (Usable) -- Not Usable otherwise

(b) Metric Usability Rating

Figure 3.6 Metric Usability Questionnaire and Rating for Base Metrics

The values of metric usability factors are formed into a vector and evaluated to determine the metric usability result. Factor values are evaluated in the order of criticality of the attributes (1 being the most critical): 1) metric identity, 2) data existence, 3) data verifiability, and 4) data dependability. The regarding values of the vector should be at least [F, F, L, L] for a base metric to be usable (vector values of [F, F, L, P], for example, leads to a result of "not usable"). For a derived metric, vector values are evaluated together with the values of metric usability factors 3 and 4 of the base metrics that make up the derived metric. Metric usability factors of 3 and 4 of the base metrics should have a value of either F or L. A value of P or N for these attributes of a base metric leads to a result of "not usable" even if usability factor values of the derived metric satisfy [F, F, L, L]. Here we should not that while coding metric usability factors 3 and 4 of the base metrics for evaluation of usability of the derived metric; we take the lowest ordinal value.

		Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Metric Name:		F : Indicatiors of the atribute are fully satisfied (%86-100)		
Conceptual Definition:		L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed On:		P : Indicatiors of the atribute are largely satisfied (%16-50)		
Assessed By:		N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes		**Answers**	**Rating**	Expected Answers
	Indicators			
Metric Identity		**MUF-1**	**F**	
Q1	What is the the metric formula? (please refer to related base metrics)			
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)			Ratio, Absolute
Q3	What is the unit of the metric data?			
Q4	What is the type of the metric data? (integer, real, etc.)			
Q5	What is the range of the metric data?			
Data Existence		**MUF-2**	**F**	
Q6	Is metric data existent?			Available > 10
Q7	What is the amount of overall observations?			
Q8	What is the amount of missing data points?			
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)			
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)			
Data Verifiability		**MUF-3**	**F**	
Q11	How is metric data calculated? (by a tool, manually, etc.)			
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)			Yes
Q13	Is all metric data calculated according to metric formula?			Yes
Q14	Where is metric data stored? (in a file, database, etc.)			
Q15	Is all metric data stored in the same place? (in a file, database, etc.)			Yes
Data Dependability		**MUF-4**	**F**	
Q16	Is metric data stored precisely?			Yes
Q17	Is metric data stored for a specific purpose?			Yes
Q18	Is the purpose of metric data storage known by process performers?			Yes
Q19	Is metric data analyzed and reported?			Yes
Q20	Is metric data analysis results communicated to process performers?			Yes
Q21	Is metric data analysis results communicated to management?			Yes
Q22	Is metric data analysis results used as a basis for decision making?			Yes
Data Normalizability				
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)			
Data Integrability				
Q24	Is metric data integrable at project level?			
Q25	Is metric data integrable at organization level?			

(a) Metric Usability Questionnaire

Metric Name:		
Conceptual Definition:		
Assessed On:		
Assessed By:		
Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	F	L or F
MUF-3&4 for base metric-1	F	L or F
MUF-3&4 for base metric-2	F	L or F
MUF-3&4 for base metric-n	F	L or F
Metric Usability Result	**F**	L or F (Usable) -- Not Usable otherwise

(b) Metric Usability Rating

Figure 3.7 Metric Usability Questionnaire and Rating for Derived Metrics

For example, assume that we are evaluating the usability of "defect density" derived metric and rate the attribute values as [F, F, F, L]. If the values of metric usability factors 3 and 4 of base metric "number of defects" are [F, L], we code the factors as "L" (the lowest of [F, L]) as basis for evaluating usability of "defect density". Similarly, if the values of metric usability factors 3 and 4 of base metric "product size" are [L, L], we code the factors as "L" again (the lowest of [L, L]). Then, since the metric usability factors of "defect density" are rated as [F, F, F, L] and the usability ratings for factors 3 and 4 for both base metrics are "L", we conclude that "defect density" derived metric is *usable* for statistical analysis. However, if the value of metric usability factor 3 or 4 was P for any of the base metrics, "defect density" would *not be usable* for statistical analysis.

Process Execution Questionnaire is a form used to investigate the external factors that might affect a process execution so that assignable causes exist. External factors are questioned in terms of changes in process performers, process environments, and other factors if any (Figure 3.8). While working retrospectively on existing process data, this form is used to understand the assignable causes for a process execution if it led to an out-of-control point. In a prospective study, however, the form is filled for each instance of the process in execution to identify the external factors that might be a potential assignable cause.

Process Name:		Recorded On:	
Process Execution No:		Recorded By:	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?		
Q2	Are process performers experienced in their roles in the process?		
Q3	Are process performers differed per role basis during execution of the process?		
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?		
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)		
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)		
Q7	Has there been a recent change in funding and resources allocated for the process?		
Q8	Has the process been tailored for this specific execution?		
OTHER FACTORS (Please list if any)			

Figure 3.8 Process Execution Questionnaire

3.4. Relation of the Model with the CMMI's Measurement Practices

How does SPC-AM differ from (or show similarity to) the measurement practices in the software industry, specifically in CMMI? We chose CMMI here, because it describes all related key practices for measurement, as independent from a specific process or tool. The traceability is provided in Table 3.2.

CMMI recommends organizational maturity or process capability levels as means of satisfying process consistency. If maturity or capability level is above 3, then we have more confidence that the processes are executed in accordance to their definitions. Level 4 requires use of statistical techniques such as SPC to ensure process stability. These requirements are not very easy to understand and interpret for emergent software organizations. CMMI provides examples of process attributes such as outputs, activities, roles, and tools and techniques; however they are all static. CMMI does not provide any guidance on how to judge process consistency based on the values of these attributes of process executions. It assumes that after recommended practices are defined, they will just be followed. However, this is not always the case. Practices change in execution according to contextual requirements that generally cannot be followed or understood. Accordingly, we developed assets to record the contextual information of a process execution and to identify rational samples based on this information. Our assumption was that process executions would be consistent within each rational sample for which we can then try statistical process control.

Measurement and Analysis process area of CMMI includes a number of recommended practices, some of which have one-to-one correspondence with our metric usability attributes. This correspondence ensured us that we have the same common understanding with the industry on what should be done to create usable metrics and data. The practices of Measurement and Analysis process area describe the things that should be done, but do not investigate what is being done. Accordingly, we developed questionnaires for base and derived metrics to understand the characteristics of metric data and related measurement practices in execution. The questionnaires included questions that investigate CMMI's measurement and analysis practices and also other usability attributes such as metric identity, data existence, and data normalizability.

Table 3.2 Model's Traceability to CMMI's Measurement Practices

SPC-AM Components	**CMMI's Measurement Practices**
PROCESS CONSISTENCY	Process consistency can be judged by looking at the capability or maturity level. There is no expected process at maturity level 1. At capability level 1, the base practices of a process are checked to see whether they are applied during process executions, and generate expected outputs. At level 2, each project has its own processes and outputs defined in the project plan, and executions of a process are monitored against the plan via process and product quality assurance activities. This makes consistent execution of a process likely within a specific project (but not organization-wide). At level 3, processes are defined organization-wide, and it is more likely that processes are executed consistently within the organization. At level 4, the processes executed by the projects are monitored against organizational performance baselines. Process consistency is ensured by statistical methods. Level 5 guides continuous process improvement, based on consistently executed processes.
Inputs	Not defined explicitly.
Outputs	Typical outputs are defined for each process area.
Activities	Goals and practices are defined for each process area, without sequencing.
Roles	Typical roles are defined for each process area.
Tools & Techniques	Tools are not defined, techniques are exemplified where appropriate.
METRIC USABILITY	Measurement and Analysis process area defines base practices to have meaningful measurement results. It is expected that when a project or organization applies these practices, it will have metrics usable for statistical analyses.
Metric Identity	Requires operational definitions of measures, but does not define what such a definition should include.
Data Existence	Requires data collection in accordance to specifications. It does not state the amount of data for analysis.
Data Verifiability	Requires definition of data collection and storage procedures. There is full correspondence here.
Data Dependability	Requires definition of data analysis and reporting procedures. There is full correspondence here.
Data Normalizability	It does not provide guidance.
Data Integrability	Organizational Process Definition process area requires establishment of a measurement repository to integrate data. It does not provide explicit guidance.

CHAPTER 4

APPLICATION OF THE MODEL

We applied the Assessment Model for Statistical Process Control in different contexts to observe its usability and to gather feedback to refine the model. During these applications we intended to investigate the answers to the question "Can an emergent organization apply SPC techniques following the guidelines proposed by our model and benefit from the results?". Our focus was specifically on working with emergent organizations or in emergent contexts because there are some well-defined frameworks to follow for large, process conscious organizations. The following sections describe the fundamentals, design, and details related to the applications.

4.1. Fundamentals of Design

There are four basic types of case study design, based on a 2 x 2 matrix of the following attributes [44]: 1) Single-case vs. multiple-case, and 2) Holistic vs. embedded. Before mentioning the type of the design that we chose for our applications, we will summarize these attributes:

- *Single-case vs. multiple-case designs*: The single-case study is appropriate under a number of circumstances such that when the case represents the "critical case" in testing a well-formulated theory, a "unique" case in which the phenomena is so rare that it is worth to analyze, or a "typical" case where the objective is to capture the conditions of an everyday situation. It has the risk that the case may later turn out not to be the one it was thought to be at the start, and therefore should be carefully investigated. Multiple-case designs have both advantages and disadvantages in comparison to single-case designs. The evidence from multiple cases is often considered more compelling, and the study is regarded as being more robust. However, the rationale for single-case designs usually cannot be satisfied by multiple-cases. The decision to undertake multiple-case studies should not be taken lightly. Every case should serve a specific purpose within the overall scope of inquiry. Here a major insight is to follow the "replication" logic which is analogous to that used in multiple experiments. Each case must be carefully selected so that it either predicts similar results or predicts contrasting results but for predictable reasons.

- *Holistic vs. embedded designs*: When the same case study may involve more than one unit of analysis, then it should have an embedded design. This occurs when, within a single case, attention is also given to a subunit or subunits. A pitfall of the embedded design takes place when the case study focuses only on the subunit level and fails to return to the larger unit of analysis. In contrast, if the study examines only the global nature of an organization or a program, a holistic design is preferred. Holistic design is advantageous when no logical subunits can be identified or when the relevant theory underlying the case study is itself of a holistic nature. A typical problem with the holistic design is that the entire case study may be conducted at an abstract level, lacking any data in detail.

We designed our applications as a multiple-case study, and identified our unit of analysis as "process-metric" pair. Since we expected that every case would include more than one unit of analysis, we decided the multiple-case design to be embedded. A process could be assessed with more than one metrics as well as more than one process (and related metrics) could be assessed in the same context. The structure of embedded, multiple-case design used for our applications is shown in Figure 4.1.

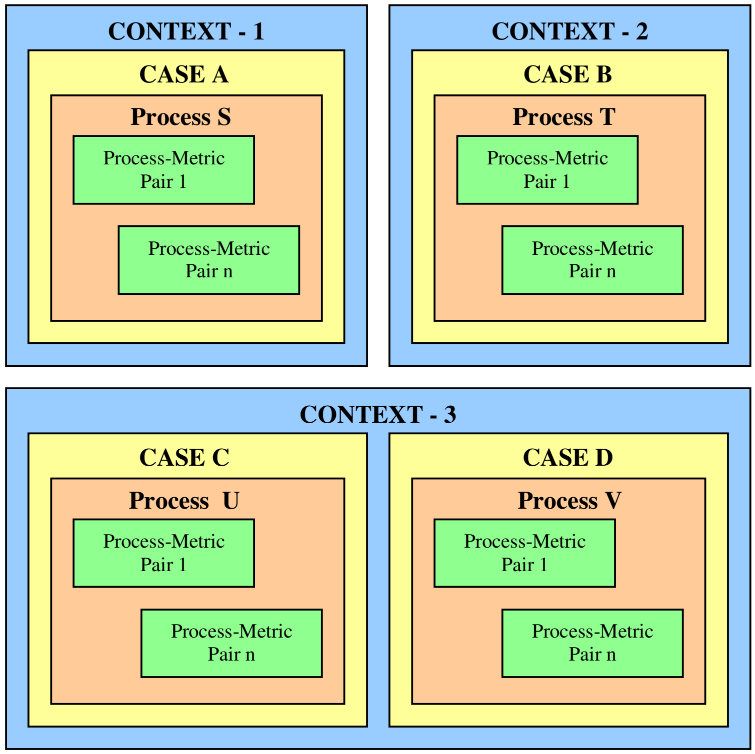


Figure 4.1 The Structure of Embedded, Multiple-Case Design

Before selecting the cases of our multiple-case study, we identified a number of criteria to consider while selecting the cases among nominations:

- Organizational size and maturity: we intended to select emergent organizations or contexts;
- Historical process execution: at least 20-25 metric data points are required;
- Accessibility of performers of historical process executions: performers will be interviewed during the assessment of process consistency;
- If there is no historical data, ability of the process to generate 20-25 metric data points in the near future;
- If there is no historical data, permission to join to future process executions as an observer to assess process consistency;
- Availability of process performers to participate in the assessment.

4.2. Multiple-Case Study Design

We used the criteria listed in the previous section, and specifically paid attention to the issue that the processes within the organization or context we would choose had not yet been fully institutionalized. This was to test our proposal that the model is expected to work for emergent organizations. If we speak in CMM or CMMI terminology, we would prefer the organizations having maturity level below 3, or the contexts (e.g. a specific project) whose practices are not yet implemented organization-wide if the organization has a maturity level 3 or above. Accordingly, we identified three organizations to perform our cases per our unit of analysis -- "process-metric" pair. In the third organization listed below, we planned to assess more than one process and to perform further evaluations on their results for checking interrelations between process performances.

- Context-1 (organization X): This is a project office of a large government research agency, which develops both systems and software for 7 years. The office usually undertakes projects to develop software for military systems, and has 18 staff including the project manager. It documented the procedures that the staff applies at a high level, and has ISO 9001 [45] certificate as related with the organizational body. The project office has been pursuing process improvement studies to achieve CMMI L3 for 20 months. It did not have a specific measurement process, but was collecting data over the tools that the staff uses in their projects.
- Context-2 (organization Y): This is a system and software development organization which has 15 years of experience in the sector and supplies products for Turkish Armed Forces with its 45-staff development team. It already has ISO 9001 [45] and AQAP-150 [46] certificates, and has been pursuing process improvement studies to achieve CMMI L3 certification for 16 months. The company did not have a specific measurement process, but was obeying policies for analyzing the data and reporting the results to high-level management. The results reported to the management were not systematically used for decision-making purposes.

- Context-3 (organization Z): This is a system and software development organization which has 16 years of experience in the sector and develops military and avionics projects with its 75-staff development team. It already has ISO 9001 [45], AQAP-150 [46], and CMM [22] L3 certificates, and is currently targeting to achieve CMMI [2] L5. Although the processes in the organization are largely defined, we worked in the context of an avionics project whose test practices have not yet been institutionalized (therefore, we did not violate our emergent case requirement).

We worked on task management; review; test design, test procedure development, and test development peer-review processes; in above contexts, respectively. We list the units of analysis for each case below. The plans for conducting the case studies, and work breakdown structure that guides each case, are provided in Appendix B.

- Case Study-A (in Context-1): We investigated utilization of “estimation capability” and “effort variance” metrics of *task management* process.
- Case Study-B (in Context-2): We worked on “non-conformance detection efficiency”, “non-conformance resolution efficiency”, “review open period”, and “review open period with respect to non-conformances” metrics of *review* process.
- Case Study-C (in Context-3): We investigated the utilization of “productivity” and “percent of internal review effort” metrics of *test design* process.
- Case Study-D (in Context-3): We investigated the utilization of “productivity” and “percent of internal review effort” metrics of *test procedure development* process.
- Case Study-E (in Context-3): We investigated the utilization of “action item density”, “action item detection efficiency”, “action item resolution efficiency” metrics of *test development peer-review* process.

The details related to each case are explained in the following sections. During the cases, we demonstrated practical evidence on the utilization of SPC via control charts. We used Minitab Statistical Software [47] to generate the charts, and applied the following tests to detect the out-of-control points:

- 1 point > 3 standard deviations from center line
- 9 points in a row on same side of center line
- 2 out of 3 points > 2 standard deviations from center line (same side)
- 4 out of 5 points > 1 standard deviation from center line (same side)

4.3. Context-1 (Case Study-A)

The task management process workflow had been defined on a change/ configuration management tool at the end of the year 2002, and had been executed for a project for about 16 months until the end of the project in March 2004. The states of the task management process, which are defined as a workflow on the tool, are shown in Figure 4.2.

Every task of the project was entered into the tool by a task assigner (Project Manager or Team Leader) with the fields of task name, responsibility, estimated start date, and estimated finish date. The responsible person then started to work on the task by recording the field for actual start date into the tool. When the task was finished, the responsible person entered actual finish date into the tool, and the task was closed after verified by the task assigner.

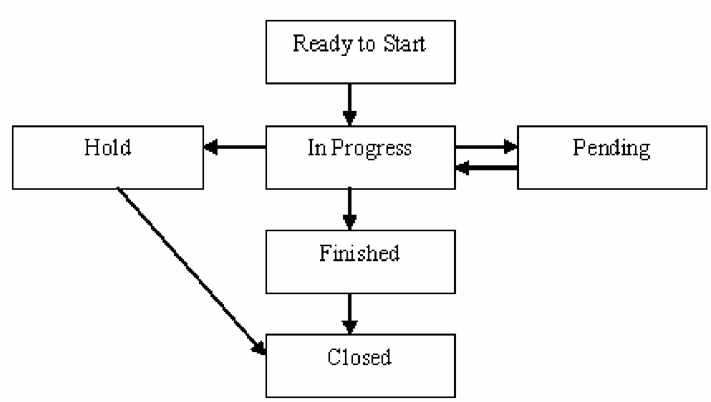


Figure 4.2 Task Management Process States

While performing the case study, we spent 6 person-days for applying the approach, performing the analyses, and interpreting the results. We started to work on existing task management data of 92 data points which were collected for the project during 16 months. The complete set of assets produced during the case together with the control charts are provided in Appendix C. We explain the case steps below over the representative assets.

Since the study was retrospective, we identified process attributes of task management process executions by inspecting the records entered into the tool and consulting the process performers. Before constructing the matrix, we sampled 4 task records from the set of 91 and filled the process execution record, as shown in Figure 4.3, for each.

Process Name:	Task Management	Recorded On:	7.June.2005
Process Execution No:	25	Recorded By:	AT

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	Task request	

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Software code	"Data Management paketine Transaction Manager class'inin tanimlanmasi"

3. Activities (in sequence): Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Enter task request	
2	Implement task request	
3	Verify implementation	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Task assigner	
2	Task implementer	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
1	Starteam	

Figure 4.3 Process Execution Record Used to Sample Task Management Process Executions

We recorded the values of process attributes on sample process execution records into the process similarity matrix, and checked out these values against 91 executions. The appearance of process similarity matrix for the first 20 executions was as in Figure 4.4.

			Process Executions																			
Process Attributes			PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20
1	**Inputs**																					
	1.1	Task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2	**Outputs**																					
	2.1	Document	•								•	•	•									
	2.2	Software code		•	•										•	•	•					
	2.3	Analysis knowledge					•	•										•		•	•	•
	2.4	Design							•	•									•			
	2.5	Research knowledge				•																
	2.6	Unclassified output												•								
3	**Activities**																					
	3.1	Enter task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	3.2	Implement task request	•	•	•	•		•		•	•	•	•	•			•	•	•	•	•	•
	3.3	Verify task request				•				•							•			•		•
4	**Roles**																					
	4.1	Task assigner	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
	4.2	Task implementer	•	•	•	•		•	•	•	•	•	•	•			•	•	•	•	•	•
5	**Tools and Techniques**																					
	5.1	Starteam	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•

Figure 4.4 Process Similarity Matrix for (the first 20) Task Management Process Executions

After finalizing the process similarity matrix, we analyzed it for similarity and differences in process executions. We identified 4 process clusters labeled from A through D as shown in Figure 4.5, by observing the similarities between process executions. We primarily searched for similarities in process attributes other than the outputs. For example, we put process executions 13 and 5 in the same cluster A, rather than putting process execution 13 and 31 in the same cluster, though the cluster distance for both pairs was 1 as it can be seen in Figure 4.5. This was because we thought the performance of task management process was less dependent on the type of output produced by the task, since the process of producing each output type (e.g., producing software code) was indeed different from the process of task management. Therefore, we used the differences in values of output attribute to categorize each process cluster into its sub-clusters (numbered from 1 to 6 with respect to output value type) for if detailed analysis would be needed for a process cluster. The sub-cluster types were as follows: 1) Document, 2) Software code, 3) Analysis Knowledge, 4) Design, 5) Research Knowledge, and 6) Unclassified output (admin, test, etc.).

Process Attributes	PE13	PE5	PE31	PE33	PE7	PE28	PE1	PE2	PE16	PE17	PE22	PE12	PE79	PE15	PE20	PE8	PE4	PE40
1 Inputs																		
1.1 Task request	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2 Outputs																		
2.1 Document							o						o					
2.2 Software code	o		o					o						o				
2.3 Analysis knowledge		o		o					o						o			
2.4 Design					o					o						o		
2.5 Research knowledge											o						o	
2.6 Unclassified output						o						o						o
3 Activities																		
3.1 Enter task request	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.2 Implement task request							o	o	o	o	o	o	o	o	o	o	o	o
3.3 Verify task request													o	o	o	o	o	o
4 Roles																		
4.1 Task assigner	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4.2 Task implementer			o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
5 Tools and Techniques																		
5.1 Starteam	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Process Cluster	A	A	B	B	B	B	C	C	C	C	C	C	D	D	D	D	D	D
Process Sub-cluster	2	3	2	3	4	6	1	2	3	4	5	6	1	2	3	4	5	6

Figure 4.5 Initial Process Clusters and Sub-clusters for Task Management Process

After we identified initial process clusters, we worked on process metrics to evaluate their usability for statistical analysis. We identified estimated start date, estimated finish date, actual start date, and actual finish date as base metrics of the task management process. These were the metrics for which data was available on the tool. From the base metrics, we identified estimated effort, actual effort, effort estimation capability, and effort variance as derived metrics of the review process. We could convert task duration (e.g. 3 days) to task effort (e.g. 3 man-days) since every task had been assigned a single responsible. Task management base and derived metrics are shown in Figure 4.6. The arrows show the relationships between the base metrics at upper side to the derived metrics at lower side.

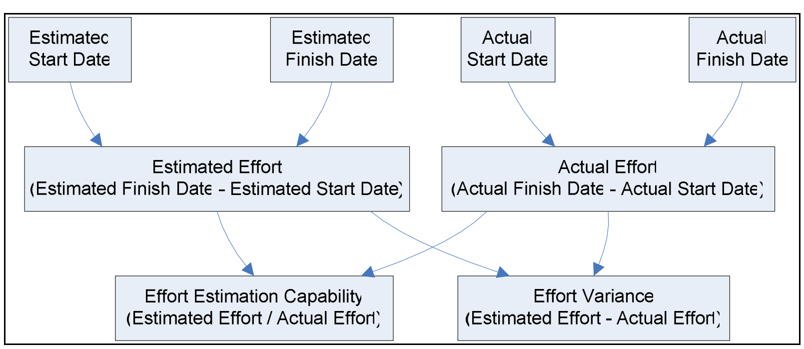


Figure 4.6 Task Management Base and Derived Metrics

We filled Metric Usability Questionnaire for each base metric shown in Figure 4.6 (completed questionnaires are provided in Appendix-C). By evaluating the answers in the questionnaires, we had judgments on the usability of base metrics first. During evaluation, we primarily paid attention to scale type and data existence requirements. There were missing data points for actual start date and actual finish date, but we thought the remaining data points could be used for analysis since data verifiability and data dependability were not violated. None of the base metrics was suitable for use on control charts, since their scale was interval. Both estimated effort and actual effort derived metrics were of ratio scale, and could be used for SPC. Similarly, effort estimation capability and effort variance were also of ratio scale, and judged as usable for statistical analysis.

We reviewed process data and used the results from process similarity assessment and metric usability evaluation to finalize process clusters and metrics prior to control charting. When we looked at data, we saw that actual start date and actual finish date fields were empty for process cluster A and that actual finish date field was empty for process cluster B. Since our derived metrics were made up of these values, we excluded process clusters A and B from our study.

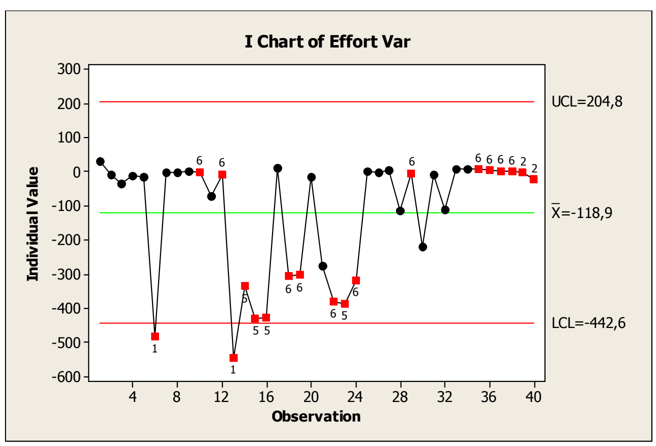
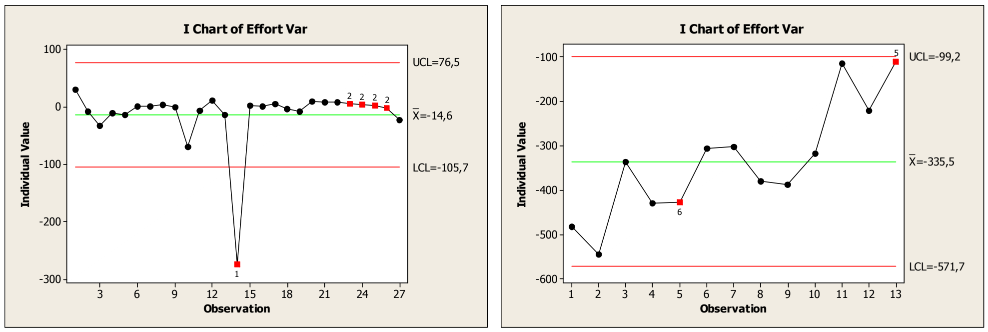


Figure 4.7 Individuals Chart for Combined Data (Clusters C and D) of Effort Variance

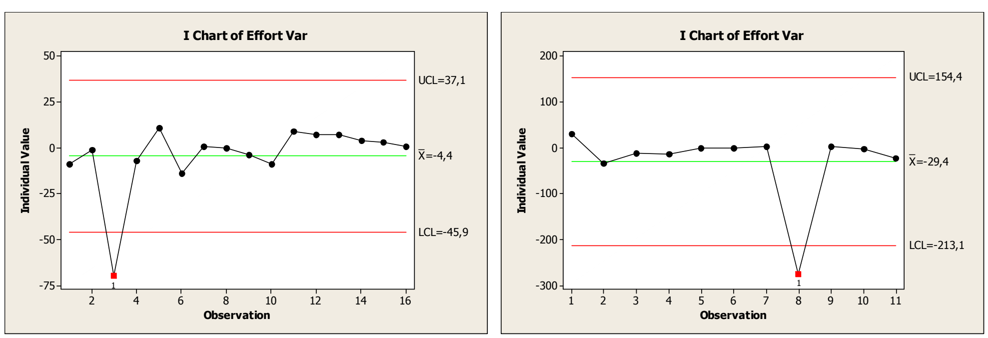
We first charted combined data of process clusters C and D, to see the current status of task management process with respect to derived metrics of estimation capability and effort variance. We applied variables charts for individuals of task management data. The chart for combined data for effort variance is provided in Figure 4.7. We then depicted task management data on control charts for process clusters C and D separately, for both derived metrics. The charts for clusters C and D for effort variance are shown in Figure 4.8.



(a) Effort variance of cluster C (b) Effort variance of cluster D

Figure 4.8 Individuals Charts for Effort Variance

From the figure above we saw that cluster C had five and cluster D had two out-of-control points with respect effort variance metric. Before investigating the assignable causes, we wanted to detail our analysis a step further for cluster C by categorizing data with respect to its sub-clusters. The number of data points for cluster C was not so much, and we observed that we could at least chart the data for output type 2 (software code) separately.



(a) Effort variance of sub-cluster C-1 (b) Effort variance of sub-cluster C-2

Figure 4.9 Individuals Charts for Effort Variance

Accordingly we identified two sub-clusters of process cluster C: C-1) Process executions having output value type of software code, and C-2) Process executions having output value type other than software code. The charts for process sub-clusters C-1 and C-2 for effort variance derived metric are given in Figure 4.9. We could not perform such an analysis for cluster-D since the number of data points was so few for sub-clustering.

We applied the same sub-clustering on process cluster C for estimation capability. Initial results obtained from control charts are summarized in Table 4.1.

Table 4.1 Initial Results from Charted Data for Task Management Process Derived Metrics

Process Cluster	**Derived Metric**	
	Estimation Capability	***Effort Variance***
Overall	Many OCPs	Many OCPs
C-1	Under control	1 OCP
C-2	1 OCP	1 OCP
D	Under control	2 OCPs

* OCP: Out-of-Control Point

We conducted an interview with project team leader in order to understand any reasons for the assignable causes. The reasons were investigated by filling process execution questionnaire for each out-of-control point reported by the table above. Process execution questionnaire for process execution-1 is shown in Figure 4.10.

Process Name: Task Management		Recorded On: 12.June.2005	
Process Execution No: 1		Recorded By: AT	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The work plan changed, and task assignment was not updated.		

Figure 4.10 Process Execution Questionnaire for Task Mgt. Process Execution # 1

Detailed findings of the reasons for out-of-control points as detected by process execution questionnaires are given in Table 4.2.

Table 4.2 Assignable Causes for Out-of-Control Points of Task Mgt. Process

Derived Metric	**Process (Sub)Cluster**	**Out-of-Control Point(s)**	**Assignable Cause**
Estimation Capability	C-1	None	Not applicable
	C-2	1	The work plan changed, and task assignment was not updated on the tool.
	D	None	Not applicable
Effort Variance	C-1	1	The work plan changed, and task assignment was not updated on the tool.
	C-2	1	Task closure was forgotten and performed late.
	D	2	Task closure was forgotten and performed late.

Based on the knowledge obtained during the interview, we re-charted the data by excluding the out-of-control points, all having assignable causes. Final results showed that process sub-clusters C-1 and C-2 and process cluster D were under control with respect to estimation capability and effort variance derived metrics.

Findings from the study:

Our model suggested that both estimation capability and effort variance derived metrics could be used for statistical analysis; and process clusters identified by process similarity assessment were detected as under control with respect to both derived metrics, at the end of the case study implementation. Due to insufficient number of data points, we could not appropriately sub-cluster the process cluster C with respect to output product type as suggested by process similarity assessment. Still, the two sub-clusters that we identified (tasks performed to generate software code vs. tasks performed to generate other work products) enabled us to validate the model's suggestions. Here we should note that the purpose for which control charts are to be used is a primary factor to determine the sub-clusters. For example, if one wanted to track effort variance of the tasks by which research knowledge was investigated (sub-cluster number C-5 in Figure 4.5), our case-study implementation would be insufficient, since we could not chart the data for sub-cluster C-5 due to lack of data points.

Primary observation from the case study was that the ranges of estimation capability were wide for process cluster C (within [-2,55 ; 5,17] with a mean value of 1,31) in which the tasks were forgotten to close frequently. Cluster-C (and its sub-clusters) included process executions for which verification activity had not been performed, and therefore we could not be sure on the validity of actual effort spent for these tasks. Cluster D, however, included verified tasks, and effort estimation capability varied within [-0,045; 0,088] with a mean value of 0,0215. Accordingly, we recommended to the team leader to check open status of tasks regularly and to perform their closures on time as possible. We also suggested that task assignments should be updated as the overall project plan changes to maintain consistency between the work assignments.

The ranges for effort variance were wide for both cluster C and D, the later showing a higher variance. Cluster C2, including task assignments for software code development, had the smallest variance (within [-20,58 ; 20,45] days with a mean value of -0,07, after removing the only out-of-control point). We attributed high variance to different sizes of work performed by the tasks. On the other hand, the control chart for effort variance of cluster D (please refer to Appendix C, Figure C.27) showed that the difference between estimated and actual effort values decreased in time, meaning an improvement in the tasks management process.

After case study implementation, we completed a process attributes description for each process cluster to demonstrate different process versions in execution. The descriptions are provided in Appendix C (please refer to Figure C.15, Figure C.16, and Figure C.17).

Reflections for improvement of the model:

The task of identifying clusters was tough due to lack of established rules and steps, and this experience helped us to formalize the identification process. After the case study, we defined the term “cluster distance” and proposed that the clusters and possible merges among them can be identified by forming a “cluster distance triangle”, as described in Chapter 3.

The judgment of metric usability (just as the judgment of similarity) was due to lack of established rules, and we observed that we need a formal evaluation process for metric usability. We thought the evaluation might be easier if we could question base and derived metrics separately. After the case study, we developed separate questionnaires for base and derived metrics, and set the rules for rating the answers in the questionnaires to evaluate metric usability. Initial versions of metric usability questionnaires are provided in the next subsection over the implementation of case study B, and the final status of the usability evaluation process is explained in Chapter 3.

4.4. Context-2 (Case Study-B)

The review process had been documented at the beginning of the year 2004, and has been in use by the staff since then for reviewing system and software development documents as well as software code. The review process definition basically included planning, review, product update, and closure activities, and directed the usage of process assets like review form and review report. The review process as defined in the company-specific procedure is given in Figure 4.11. Every review started with an announcement of the review, and completed when the product was accepted without any non-conformances. In other words, if a number of non-conformances were detected, the review was not closed until all of them were removed from the product. The review data was recorded on a review form when non-compliances were detected in the product, and on a review report at the end of the review. The review form included the fields for review date, product description (type, version, and configuration item type that the product belongs to), related project, and review effort as well as a list for non-conformances. The review report included the fields for review date, closure date, participants, product description, related project, review type (internal or joint), review result, number of non-conformances detected, number of non-conformances accepted, review effort, and non-conformance resolution effort if any. The results and data from completed reviews were reported to the Quality Manager of the company every month.

While performing the case study, we spent 14 person-days for gathering and translating review data, applying the approach, performing the analyses, and interpreting the results. We worked on existing review process data of 200 data points which were collected during two years. We translated the review data to a form that is appropriate for comparison among different projects and products. We extracted 4 reviews from the set because the data was missing or not properly recorded, and dealt with the data from remaining 196 reviews. The complete set of assets produced during the case together with the control charts are provided in Appendix D. We explain the case steps below over the representative assets.

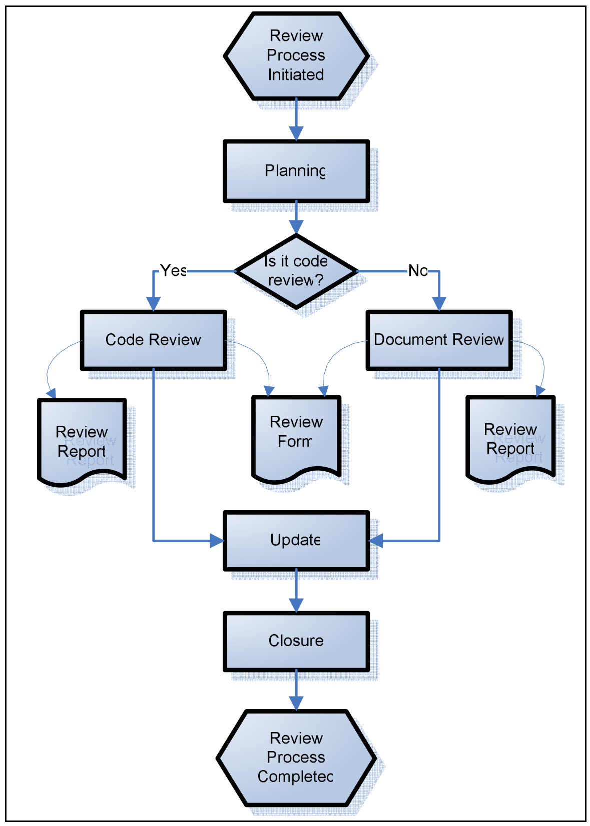


Figure 4.11 Review Process as Defined in the Company-Specific Procedure

Since the study was retrospective, we identified process attributes of review process executions by inspecting review process outputs (review forms and review reports) and consulting the Quality Assurance Expert participated in the reviews. Quality Assurance Expert is a staff of the Quality Department and is responsible for coordinating and following quality assurance and configuration management activities per project basis. Before constructing the matrix, we sampled 5 reviews from the set of 196 and filled the process execution record, as shown in Figure 4.12, for each.

Process Execution Record
(Internal Attributes)

Process Name:	Review	Recorded On:	27/09/2005
Process Execution No:	10	Recorded By:	A.Tarhan

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	SDD	Software Design Document

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Review Form	
2	Review Report	

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Planning	
2	Review	
3	Update after meeting	
4	Closure	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
	NOT RECORDED	

Figure 4.12 Process Execution Record Used to Sample Review Process Executions

			Process Executions																				
Process Attributes			PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	....
1	**Inputs**																						
	1.1	Product type to review	SPS	SSS, SSDD	SRS	STP	STD	SDD	DB DD	YST, SPS	SRS	STP	STD	SDD	STP	SDP	STD	SRS	YGD	YGD	AKD 1	AKD 1	
2	**Outputs**																						
	2.1	Review form	o		o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	
	2.2	Review report	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
	2.3	Code review report																					
3	**Activities**																						
	3.1	Planning	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
	3.2	Review	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
	3.3	Update during meeting																					
	3.4	Update after meeting	o		o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	
	3.5	Closure	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
4	**Roles**																						
	4.1	Project Manager	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
	4.2	QA Expert	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
	4.3	CM Specialist	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	
	4.4	Customer																					
5	**Tools and Techniques**																						
		NOT RECORDED																					

Figure 4.13 Process Similarity Matrix for (the first 20) Review Process Executions

The information on process execution records provided us typical values of process attributes, and formed an initial base to create the process similarity matrix. After we recorded these typical values of process attributes on the matrix, we checked out these attributes for all 196 executions. The appearance of process similarity matrix for the first 20 executions was as in Figure 4.13. Once the matrix was completed, we reviewed the attribute values for potential abnormalities that might affect our analysis. We detected that more than one product was input for review in some executions (e.g., 2 and 8 in Figure 4.13). We excluded 4 such executions from the set, and came up to 192 process executions as a result.

After finalizing the process similarity matrix, we analyzed it for similarity and differences in process executions. We identified 9 process clusters labeled from A through I, as shown in Figure 4.14, none of which were the same in terms of attribute values.

2	**Outputs**										
	2.1	Review form	o		o		o		o	o	
	2.2	Review report	o	o	o		o	o	o	o	o
	2.3	Code review report				o					
3	**Activities**										
	3.1	Planning	o	o	o	o	o	o	o	o	o
	3.2	Review	o	o	o	o	o	o	o	o	o
	3.3	Update during meeting					o		o		
	3.4	Update after meeting	o		o						o
	3.5	Closure	o	o	o	o	o	o	o	o	o
4	**Roles**										
	4.1	Project Manager	o	o	o	o	o	o	o	o	o
	4.2	QA Expert	o	o	o	o	o	o	o	o	o
	4.3	CM Specialist	o	o	o	o	o	o	o	o	o
	4.4	Customer			o			o	o	o	
5	**Tools and Techniques**										
		NOT RECORDED									
Process Cluster			A	B	C	D	E	F	G	H	I

A-B	2														
A-C	1	B-C	3												
A-D	4	B-D	2	C-D	5										
A-E	2	B-E	2	C-E	3	D-E	4								
A-F	3	B-F	1	C-F	2	D-F	3	E-F	3						
A-G	3	B-G	3	C-G	2	D-G	5	E-G	1	F-G	2				
A-H	2	B-H	2	C-H	1	D-H	4	E-H	2	F-H	1	G-H	1		
A-I	1	B-I	1	C-I	2	D-I	3	E-I	3	F-I	2	G-I	4	H-I	3

Cluster		Mergable To	Will Be Merged To
B	→	--	--
C	→	A	A
D	→	--	--
E	→	--	--
F	→	B	B
G	→	E	E
H	→	C,F,G	--
I	→	A,B	A

Figure 4.14 Initial Process Clusters and Cluster Distances for Review Process

Each process cluster we identified was a rational sample of the review process, and ideally we could chart the data for each cluster to see whether the process cluster was under control. When we counted the number of process executions in the clusters, we noticed that many clusters (except A and B) had few executions. We could either remove the clusters with few data from the set and continue our study with clusters A and B only, or find a way to merge the clusters with limited data to some other cluster. We chose the latter.

To identify possible merges among the clusters, we first calculated the cluster distance for each pair of clusters as shown in the upper right corner of Figure 4.14. We then recorded the mergable clusters in a table shown in the lower right corner of the figure. Some clusters were not mergable to any other cluster, and some were mergable to more than one cluster. We randomly chose to merge cluster I to cluster A. We excluded cluster H from the study due to few number of data points.

As a result of the clustering process, we ended up with the following clusters: Cluster A (including initial clusters A, C, and I); cluster B (including initial clusters B and F); cluster D; and cluster E (including initial clusters E and G). Cluster D entirely included process executions for code review. Unfortunately, the number of data points for cluster D was so few that we excluded it from the study.

After we identified initial process clusters, we worked on process metrics to evaluate their usability for statistical analysis. We identified review opening date, review closure date, number of detected nonconformances, number of accepted nonconformances, and nonconformance resolution effort as base metrics of the review process. These were the metrics for which data was available on review records. From the base metrics, we identified review open period, review open period with respect to nonconformances, nonconformance detection rate, and nonconformance resolution rate as derived metrics of the review process. Defect detection rate, defect removal rate, and defect density were the most popular metrics for review process in the literature [12][13][14][15]. Product size had not been recorded regularly so that we could not chart the data for nonconformance density metric. Instead, we later utilized software product's LOC data, which was recorded partially per month basis for year 2005, in order to rationalize nonconformance detection efficiency while evaluating process performance.

The set of review metrics that we worked on are shown in Figure 4.15. The arrows show the relationships between the base metrics at upper side to the derived metrics at lower side. The derived metrics are calculated from the base ones by the formulas described below:

- Open period: Closure date – Opening date
- Open period with respect to nonconformances: Open period / Number of accepted nonconformances
- Nonconformance detection efficiency: Number of accepted nonconformances / Review effort
- Nonconformance resolution efficiency: Number of accepted nonconformances / Nonconformance resolution effort

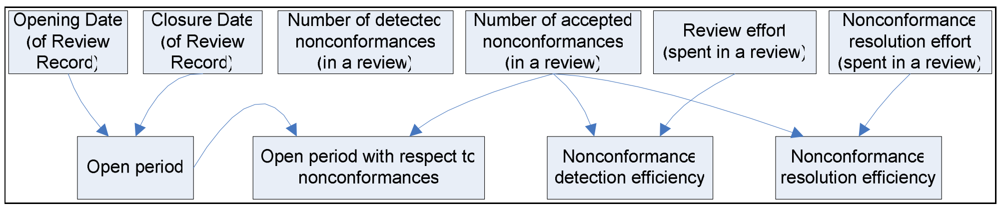


Figure 4.15 Review Base and Derived Metrics

For evaluating the usability, we used separate questionnaires for base and derived metrics. Figure 4.16 and Figure 4.17 provide examples of metric usability questionnaires and calculated metric usability indices for review effort and nonconformance detection efficiency metrics. Individual points for metric identity and data existence attributes for review effort are 1, meaning that the metric is usable at the first place. The points for data verifiability and data dependability attributes are 1 and 0.75 respectively, leading a usability index of 0.75 when multiplied. For noncompliance detection efficiency, we again have usability in the first place since individual points for metric identity and data existence attributes are both 1. Data verifiability is pointed as 1 and data dependability cannot be pointed since metric data is lately calculated by us and not by process performers during process executions. Metric usability index is calculated by multiplying the only point for data verifiability (1) by the arithmetic mean of the usability indices of regarding base metrics (both 0.75). The resulting metric usability index for nonconformance detection efficiency is therefore 0.75.

Metric Name: Review effort			**Assessed On:** 11/10/2005	
Conceptual Definition: Total effort spent for product review			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	Which entity does the metric measure?	Human resource	
	Q2	Which attribute of the entity does the metric measure?	Effort	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q4	What is the unit of the metric data?	Man-hour	
	Q5	What is the type of the metric data? (integer, real, etc.)	Real	
	Q6	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	63	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	Yes (1-58)	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the review meeting	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,75**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for QMS)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	Yes (as total review effort by projects)	1
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	Yes	1
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q31	Is metric data integrable at project level?	Yes	
	Q32	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,75**

Figure 4.16 Metric Usability Questionnaire for "Review Effort" Base Metric

Metric Name: Noncoformance detection efficiency			**Assessed On:** 11/10/2005	
Conceptual Definition: Average effort to detect a nonconformance			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	What is the the metric formula? (please refer to related base metrics)	Number of accepted nonconformances / Review effort	
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q3	What is the unit of the metric data?	Number of nonconformances per man-hour	
	Q4	What is the type of the metric data? (integer, real, etc.)	Real	
	Q5	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q6	Is metric data existent?	Yes	1
	Q7	What is the amount of overall observations?	199	
	Q8	What is the amount of missing data points?	63	
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	Yes (1-58)	
Data Verifiability				**1,00**
	Q10	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula	
	Q11	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes	1
	Q12	Is all metric data calculated according to metric formula?	Yes	1
	Q13	Where is metric data stored? (in a file, database, etc.)	In excel sheet	
	Q14	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**N/A**
	Q15	Is metric data stored precisely?	Yes	1
	Q16	Is metric data stored for a specific purpose?	Yes	1
	Q17	Is the purpose of metric data storage known by process performers?	N/A	
	Q18	Is metric data analyzed and reported?	N/A	
	Q19	Is metric data analysis results communicated to process performers?	N/A	
	Q20	Is metric data analysis results communicated to management?	N/A	
	Q21	Is metric data analysis results used as a basis for decision making?	N/A	
Data Normalizability				
	Q22	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q23	Is metric data integrable at project level?	Yes	
	Q24	Is metric data integrable at organization level?	Yes	
			Total Points	**1,00**
			MUI for no.of accepted nonconformances	**0,75**
			MUI for review effort	**0,75**
			Metric Usability Index	**0,75**

Figure 4.17 Metric Usability Questionnaire for "Nonconformance Detection Efficiency" Derived Metric

The results of usability evaluations for all review metrics in our assessment are given in Table 4.3.

Table 4.3 Usability Evaluation Results for Review Process Metrics

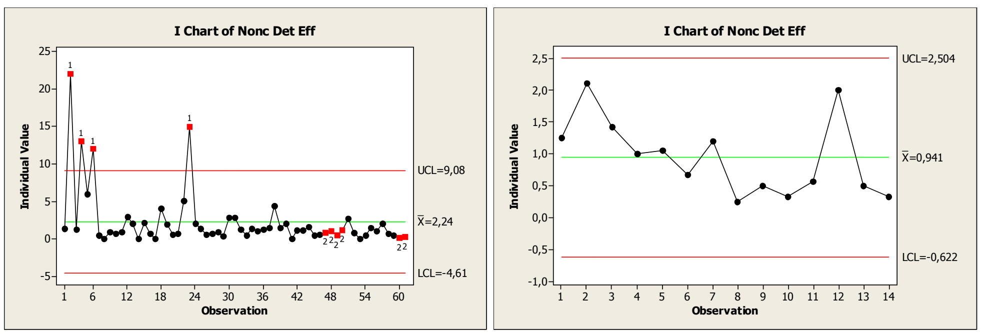
Metric	Metric Usability Index	Usability Status
Opening date	0.00	Not Usable [0.00-0.25]
Closure date	0.00	Not Usable
Number of detected nonconformances	0.50	Poorly Usable [0.26-0.50]
Number of accepted nonconformances	0.75	Largely Usable [0.51-0.75]
Rewiew effort	0.75	Largely Usable
Nonconformance resolution effort	0.75	Largely Usable
Open period	0.50	Poorly Usable
Open period with respect to nonconformances	0.58	Largely Usable
Nonconformance detection efficiency	0.75	Largely Usable
Nonconformance resolution efficiency	0.75	Largely Usable

We reviewed process data and used the results from process similarity assessment and metric usability evaluation to finalize process clusters and metrics prior to control charting. Accordingly, we performed the following:

- We intended to work with the data for derived metrics having metric usability index greater than 0.50. Open period with respect to nonconformances, nonconformance detection efficiency and nonconformance resolution efficiency were such metrics. We later included the data for open period derived metric for control charting, since it had a metric usability index of 0.50 which was very close to the lower limit for large usability. We did not intend to chart the data for any of the base metrics because they needed to be normalized for effective use.
- We noticed that process cluster B (with initial clusters B and F) included process instances in which no nonconformance was detected. It would not be meaningful to chart the data for nonconformance detection efficiency, nonconformance resolution efficiency, and open period with respect to nonconformances derived metrics in this case, since all values would be zero according to their formulas. Therefore, we excluded cluster B from the study.

As a result, we chose two clusters as basis for control charting with derived metrics: Process cluster A (including initial clusters A, C, and I) and process cluster E (including initial clusters E and G). We renamed these clusters as M and N, respectively, to distinguish them from their homonymous initial clusters. If we should describe the characteristics of these two clusters from the process view, we note that process cluster M included process executions in which product updates are performed after the review, and process cluster N included process executions in which product updates are completed within the review.

We depicted review data on control charts for process clusters M (including initial clusters A, C, and I) and N (including initial clusters E and G), and for each derived metric separately. We applied variables charts for individuals of review data. The charts for clusters M and N for nonconformance detection efficiency are shown in Figure 4.18.



(a) Control chart for cluster M (b) Control chart for cluster N

Figure 4.18 Individuals Charts for Nonconformance Detection Efficiency

Table 4.4 Sub-Clusters of M with respect to Input Product Types

Sub-cluster number	**Input product type description**	**Input product type**
1	Project plans	Project management plan, quality assurance plan, configuration management plan, subcontract management plan, system engineering management plan, qualification test plan, system test plan, software development plan, software installation plan, software test plan
2	Design documents	System design document, system/subsystem design document, interface control document, database design document, software design document
3	Analysis documents	Operational concepts document, system/subsystem specification, software requirements specification, software product specification, pre-integration requirements document
4	Other documents	Qualification test procedures, software test descriptions, system test report, software test report, software version identification document, software release identification document, arguments for data dictionary, software user document

From Figure 4.18 we saw that cluster M had many out-of-control points and cluster N was under control with respect to nonconformance detection efficiency metric. At this point we categorized review data according to input product type, and obtained the sub-clusters listed in Table 4.4 for cluster M. We then separately charted the data for each sub-cluster.

We performed similar analyses for nonconformance resolution efficiency, open period, and open period with respect to nonconformances metrics. The results obtained from control charts are summarized in Table 4.5.

Table 4.5 Initial Results from Charted Data for Review Process Derived Metrics

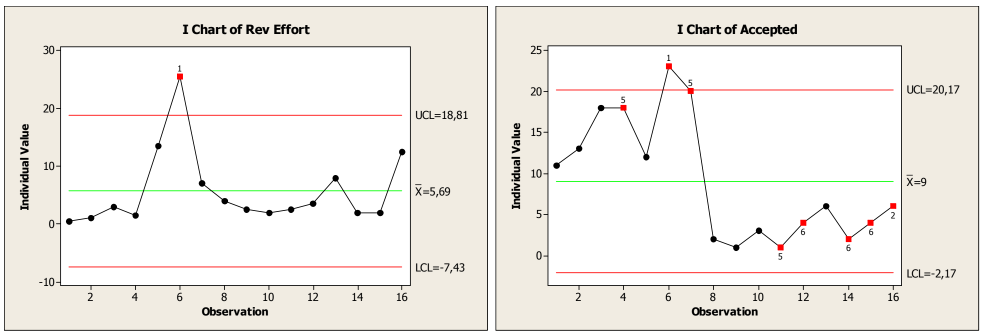
Process Cluster	**Derived Metric**			
	Nonconformance Detection Efficiency	***Nonconformance Resolution Efficiency***	***Open Period wrt. Nonconformances***	***Open Period***
Overall	Many OCPs	Many OCPs	Many OCPs	Many OCPs
M-1	Many OCPs	Under control	Many OCPs	2 OCPs
M-2	Under control	Many OCPs	Many OCPs	Many OCPs
M-3	Under control	Under control	2 OCPs	Many OCPs
M-4	1 OCP	2 OCPs	3 OCPs	2 OCPs
N	Under control	No data	2 OCPs	2 OCPs

* OCP: Out-of-Control Point

We had the following interpretations per process cluster basis from these initial results:

- Overall: Review process was not under control with respect to any derived metrics. We thought this indicated a mixture of multiple cause systems within the process, and sub-clustering supported our proposition.
- Process cluster M-1: The cluster included process executions for project plans, and had many out-of-control points with respect to nonconformance detection efficiency. Since there were different types of plans (management, quality assurance, configuration management, etc.) within the cluster and reviewing them would require considering related managerial issues, we thought the number of nonconformances might have differed for each type. We charted the data for review effort and the number of accepted nonconformances base metrics separately to check this idea (Figure 4.19). We observed that the

control chart for review effort had an out-of-control point and the control chart for number of accepted nonconformances had many, which highly supported our idea. On the other hand, the cluster had two out-of-control points for open period and many out-of-control points for open period with respect to nonconformances metric. We thought the reason was again the number of nonconformances accepted for the plans of different types.

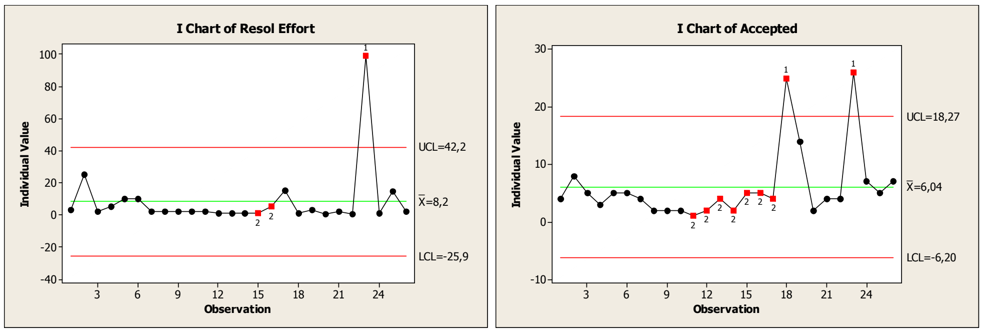


(a) for Review Effort (b) for Number of Accepted Nonconform.

Figure 4.19 Individuals Charts for Base Metrics of Nonconformance Detection Efficiency for Process Cluster M-1

- Process cluster M-2: The cluster included process executions for design documents, and was under control with respect to nonconformance detection efficiency only. We thought that the instability of data for nonconformance resolution efficiency might be due to two reasons: First, removing the nonconformances in design documents could require executing the design process and could affect the resolution effort in turn; and second, design documents were of various types so that the number of nonconformances might had differed just as in plan documents. We charted the data for nonconformance resolution effort and the number of accepted nonconformances (Figure 4.20), and observed that the control chart for nonconformance resolution effort had 3 out-of-control points and the control chart for number of accepted nonconformances had many. The findings of the charts supported our propositions. The cluster was not under control for either open period or open period with respect to nonconformances derived metrics.

We thought the latter might have been affected by the number of accepted nonconformances again, but we could not derive a clear rational for instability of open period.



(a) for Resolution Effort (b) for Number of Accepted Nonconform.

Figure 4.20 Individuals Charts for Base Metrics of Nonconformance Resolution Efficiency for Process Cluster M-2

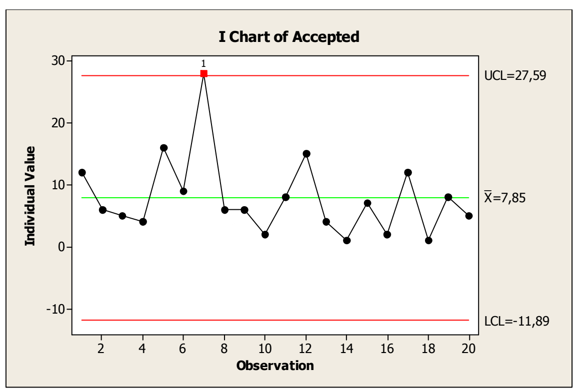


Figure 4.21 Individuals Chart for "Number of Accepted Nonconformances" Base Metric of Open Period with respect to Nonconformances for Process Cluster M-3

- Process cluster M-3: The cluster included process executions for analysis documents, and was under control for nonconformance detection efficiency and nonconformance resolution efficiency derived metrics. We thought the reason for instability of open period might be due to project characteristics, but we could not exactly identify what characteristics they were. We charted the data for number of accepted nonconformances (Figure 4.21), and observed that the chart had only one out-of control point. Therefore we thought that the reason of instability for open period with respect to nonconformances was the same as for open period.

- Process cluster M-4: The cluster included process executions for other documents, and had few number of out-of-control points with respect to all derived metrics. We attributed the few number to "other" documents which included standard, word-based work products for which we though the review process had been executed more consistently.
- Process cluster N: The cluster was under control with respect to nonconformance detection efficiency, and had two out-of-control points for each open period and open period with respect to nonconformances derived metrics. We could not foresee any reason for the out-of-control points, but we thought the reason might be the same for open period and open period with respect to nonconformances.

Process Name: Review		Recorded On: 03.11.2005	
Process Execution No: 98		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	Yes	Computer based review
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	Yes	Review was performed primarily for consistency checking (verification) - not like standard document review
OTHER FACTORS (Please list if any)			

Figure 4.22 Process Execution Questionnaire for Review Process Execution # 98

We conducted interviews with process performers in order to understand any reasons for the assignable causes. The interviews were performed in two parts. In the first part, the experiences and dynamics of process executions were investigated in free format dialogs, and notes were taken. Here the purpose was to have an understanding of the context related to process executions, and to identify any assignable cause (probably that our approach could not detected) from the performers' point of view. Process abnormalities and project specific constraints were reported by process performers in this context. In the second part, the reasons for assignable causes detected by our approach were questioned specifically. Process execution questionnaire shown in Figure 4.22 was used for this purpose.

During the interviews, the following issues were reported by process performers as potential reasons for out-of-control points:

- *Involvement of contractors in the review*: All projects had upper contractors within 2 or 3 level of subcontracting. When contractors joined in a review, generally the number of detected nonconformances increased as well as the types of nonconformances differed due to diverse points of view. Review effort also increased in some cases since it required more time to set a common understanding among the contractors not only on the meaning of nonconformances but also on some managerial aspects related to the project of concern. Update to a joint product by different parties demanded more time due to integration required for the parts of the product.

- *Project schedule*: Projects with shorter duration included frequent reviews that typically had shorter open periods. In some cases review duration was limited due to time constraints. When project duration was long, inter-review time for upgrading versions of a product increased due to longer cycle time. This led to accumulated number of nonconformances in the products. Project duration had a visible effect on the open period of a review record.

- *Product type under review*: The type of the product affected how the review was conducted in some cases. For example, the review for a software version document or a software product specification was performed primarily to verify the information in the document to the work products, on a computer screen.

The reasons for the out-of-control points were investigated in a structured manner. The underlying reason was questioned for each out-of-control point via process execution questionnaire. Detailed findings are given in Table 4.6.

Table 4.6 Assignable Causes for Out-of-Control Points of Review Process

Derived Metric	**Process Cluster**	**OCPs**	**Assignable Cause**
Nonconformance detection efficiency	M-1	Many	No specific reason for OCPs regarding PEQs. However, the OCPs belonged to a specific project which had a very tight schedule.
	M-4	1	Review performed as verification on a computer screen.
Nonconformance resolution efficiency	M-2	Many	No specific reason for OCPs regarding PEQs. However, started to record review effort in the meanwhile.
	M-4	2	Product under review was system test report. Review performed primarily for syntactical errors.
Open period with respect to nonconformances	M-1	many	No specific reason for OCPs regarding PEQs.
	M-2	many	No specific reason for OCPs regarding PEQs. However, the OCPs belonged to a specific project which had a very long schedule.
	M-3	2	No specific reason for OCPs regarding PEQs.
	M-4	3	Two subsequent versions of quality test procedures were reviewed in a single review.
	N	2	No specific reason for OCPs regarding PEQs.
Open period	M-1	2	No specific reason for OCPs regarding PEQs.
	M-2	many	No specific reason for OCPs regarding PEQs.
	M-3	many	No specific reason for OCPs regarding PEQs.
	M-4	2	Two subsequent versions of quality test procedures were reviewed in a single review.
	N	2	No specific reason for OCPs regarding PEQs.

* OCP: Out-of-Control Point, PEQ: Process Execution Questionnaire

For some clusters, no specific reason could be detected for the out-of-control points by using the process execution questionnaire, as indicated in the rightmost column of the table. According to the suggestions of our model, we expected such cases not to happen for nonconformance detection efficiency and nonconformance resolution efficiency metrics, but to happen for open period with respect to nonconformances and open period metrics, since the latter two were evaluated less usable for statistical analysis. Unfortunately there were two such cases for the former two metrics: Cluster M-1 with respect to nonconformance detection efficiency, and cluster M-2 with respect to nonconformance resolution efficiency. Although the assignable causes could not be detected by process execution questionnaires for them, we could identify underlying reasons by consulting the process performers. The out-of-control points for cluster M-1 with respect to nonconformance detection efficiency belonged to a specific project which had a very tight schedule. Starting to record review effort in the meanwhile caused a change in cluster M-2 as detected by the control chart for nonconformance resolution efficiency. However, no reason could be detected for the out-of-control points for open period with respect to nonconformances and open period metrics (except for cluster M-2 for open period with respect to nonconformances), which highly supported the suggestions of our model.

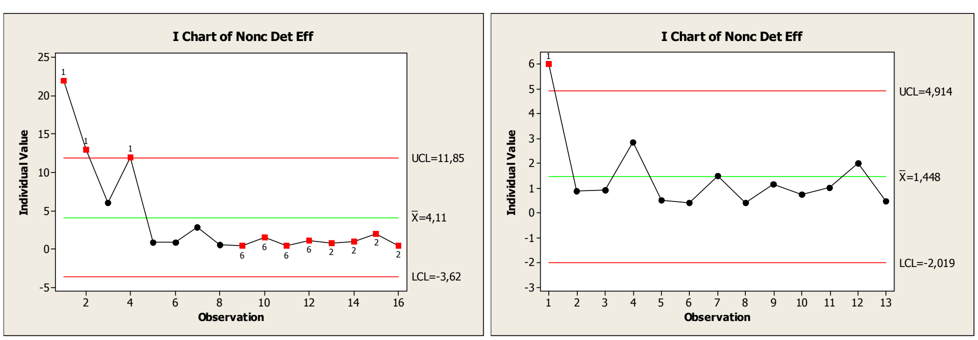
Based on the knowledge obtained during interviews, we re-charted the data by excluding the out-of-control points that had assignable causes. Final results obtained from re-charted data are summarized in Table 4.7.

Table 4.7 Final Results from Re-charted Data for Review Process Derived Metrics

Process Cluster	**Derived Metric**			
	Nonconformance Detection Efficiency	***Nonconformance Resolution Efficiency***	***Open Period wrt. Nonconformances***	***Open Period***
M-1	1 OCP	Under control	Not under control	Not under control
M-2	Under control	Splitted into two: part-1: under control part-2: 1 OCP	Not under control	Not under control
M-3	Under control	Under control	Not under control	Not under control
M-4	Under control	Under control	Under control	Under control
N	Under control	No data	Not under control	Not under control

* OCP: Out-of-Control Point

For the rest of the out-of-control points, by considering the issues reported by process performers as potential reasons for out-of-control points, we had the following decisions per derived metric basis while charting the data:



(a) Control chart for cluster M-1 (initial) (b) Control chart for cluster M-1 (final)

Figure 4.23 Re-charting of Nonconformance Detection Efficiency for Cluster M-1

- Noncompliance detection efficiency: For cluster M-1, there seemed an extraordinary situation for the first four data points as shown in Figure 4.23(a). We excluded the out-of-control points that were out of the control limits since they belonged to a specific project which had a very tight schedule. The resulting chart is shown in Figure 4.23(b). The only out-of-control point belonged to the same project as the previous ones.
- Noncompliance resolution efficiency: For cluster M-2, we divided the cluster into two parts. Both review form and review report had been updated at a time in the middle of process execution sequence to record the review effort. After splitting, part-1 included the executions for which review effort was not recorded, and part-2 included the executions in which review effort was recorded. When we put the data on charts, we saw that part-1 was under control and part-2 had an out-of-control point with an assignable cause. Review effort was not one of the base metrics of noncompliance resolution efficiency; however it was obvious, especially after re-charting, that updates in process assets led to a change in the process as well.
- Open period with respect to nonconformances: For cluster M-2, there seemed an extraordinary situation for the first six data points. We excluded four out-of-control points that were out of the control limits since they belonged to a specific project which had a very long schedule. The resulting chart included two out-of-control points for which no assignable cause could be detected.
- Open period: There was no specific reason for the out-of-control points of this metric in process clusters (except M-4), and we could not identify any action on them based on our findings.

Findings from the study:

By performing process similarity assessment based on process attributes, we could rationally sample process instances and identified two process clusters (M and N) as basis for control. We evaluated the usability of review metrics for SPC analysis, and evaluation results suggested that nonconformance detection efficiency, nonconformance resolution efficiency, and open period with respect to nonconformances metrics were largely usable for SPC analysis, the first two being more likely to succeed considering metric usability indices.

By putting process data on control charts, we identified that cluster M was out of control so that we decomposed it into four further clusters by categorizing process data with respect to product type under review. Corresponding charts were promising. We investigated any assignable causes for decreased number of out-of-control points on these charts, and removed assignable causes before re-charting the data. When re-charted the data, we observed that all process clusters were under control with respect to nonconformance detection efficiency and nonconformance resolution efficiency metrics except two out-of-control points for which we could detect assignable causes. Therefore, statistical analysis results validated the suggestions brought by our approach. Process cluster M-4 was under control with respect to all derived metrics since it included regular documents, for which the review process was affected at minimum degree by factors as project dynamics, development maturity, and etc.

Nonconformance detection efficiency and nonconformance resolution efficiency metrics were usable for SPC analysis, but would they be useful as well? Nonconformance detection efficiency metric could be an indicator of review process efficiency, but definitely not alone, since we had no idea on the defectiveness of the product under review. The defectiveness of a product is measured by defect density metric calculated by the ratio of number of nonconformances in a product to the product size. In other words, without knowing how many nonconformances included in the unit size of the product (i.e., page), we can not have a judgment on whether detecting, for example, 5 nonconformances in an hour is good or bad in terms of process efficiency. In our case, the size of the product under review was not recorded regularly, but software product's LOC data was recorded partially per month basis for year 2005. Therefore, we utilized existing LOC data to rationalize nonconformance detection efficiency for process performance. We identified reviews performed in 2005, and according to their opening dates, we recorded regarding LOC values. From the number of noncoformances accepted in these reviews and regarding software product size in LOC, we calculated nonconformance density metric by the formula “number of accepted nonconformances/KLOC” and we charted the metric data (please refer to Appendix D, control charts for defect density: Figure D.51, Figure D.52, Figure D.53). We observed that overall process had two out-of-control points, while process clusters M and N were both under control. We concluded from these findings that nonconformance detection efficiency metric can be used to judge and improve process performance since the nonconformance density metric was stable at the moment. We also noted that the company should keep recording product size to continually monitor nonconformance density for possible changes in the performance, and recommended recording the size of the product under review on process assets in each review.

Nonconformance resolution efficiency metric could be useful for planning purposes if the product was to be updated after review meeting. By looking at the number of nonconformances accepted in the review and at related control chart, process performers could estimate the effort required for resolving the nonconformances. The type of nonconformance could have a significant effect on estimating nonconformance resolution effort since syntactic errors would take less effort to fix while semantic ones would take more. Again, as a means of improvement for future use, we recommended recording the type of each nonconformance accepted in each review.

In addition to the suggestions stated above for improving process metrics and assets, identifying process clusters provided insight for improving the review process itself. By questioning process attributes of process executions, process performers could have a clearer understanding of how they changed the process for specific needs of a review. Review process flow identified after clustering is shown in Figure 4.24. We should note that although process performers mentioned joint reviews as one of the potential sources of process out-of-control points, we could not find tangible evidence to separate process flow of joint reviews from that of internal reviews.

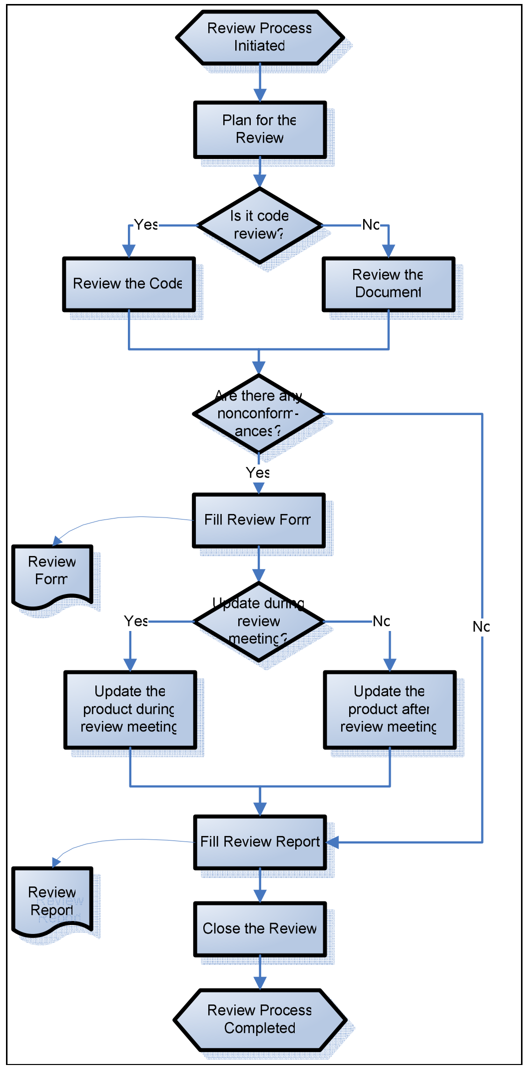


Figure 4.24 Review Process Flow after Case Study Implementation

Reflections for improvement of the model:

After the implementation of case study A, we had noted that the purpose for which control charts are to be used is a factor that should be considered while determining the sub-clusters. While performing case study B, we observed that the purpose of application also affects identification of possible merges among the process clusters. Constructing control charts for process improvement will lead to different merges and sub-clustering, when compared with the control charts for project management, for example. Therefore, we added this issue to our approach as a rule of thumb to guide identification of rational samples of a process.

Another requirement for improvement was related with evaluation procedure of metric usability. We used numeric values to weight the questions, but without any theoretical or experimental background. Metric usability attributes could be judged on ordinal scale (for example, by ranking between 1 and 5), but we utilized the weights of their questions as though they were of ratio scale. This was not valid from the perspective of measurement theory. Accordingly, we updated the evaluation process as described in Chapter 3, and evaluated the metrics' usability in ordinal scale in the following case study implementations. We utilized ISO 15504's [3] method for rating process attributes as a guide for the changes.

4.5. Context-3 (Case Studies C, D, and E)

Within the third context, we worked on test design, test script development, and test development peer review processes of an avionics project. Test development studies of the project started in September 2003 and have progressed since then for 52 test packages of three different modules, resulting with about 600 test cases at the moment that we conducted the cases. Since the studies are expected to complete within fall 2006, we could not utilize the data from entire process executions. We had to exclude 4 executions of test design and test script development processes, and 21 executions of test development peer review process while performing the cases.

We gathered data from Team Leaders of test development teams. The teams have been entering the effort for test design and test script development, the number of test cases per package, the effort for internal reviews performed during test design and test script development, the effort for peer reviews of test packages, and the number of action items detected in peer reviews, into excel sheets for the purpose of tracking. However, the Team Leaders stated that these data have not been used effectively for decision making or re-planning. Test design and test development processes have not yet been defined, but peer review process was established organization-wide. We explain details of case studies C, D, and E as a whole here, since their application and results are closely related. We spent 5 person-days for applying the approach, performing the analyses, and interpreting the results. The set of assets produced during the cases C, D, and E together with the control charts are provided in Appendices E, F, and G, respectively.

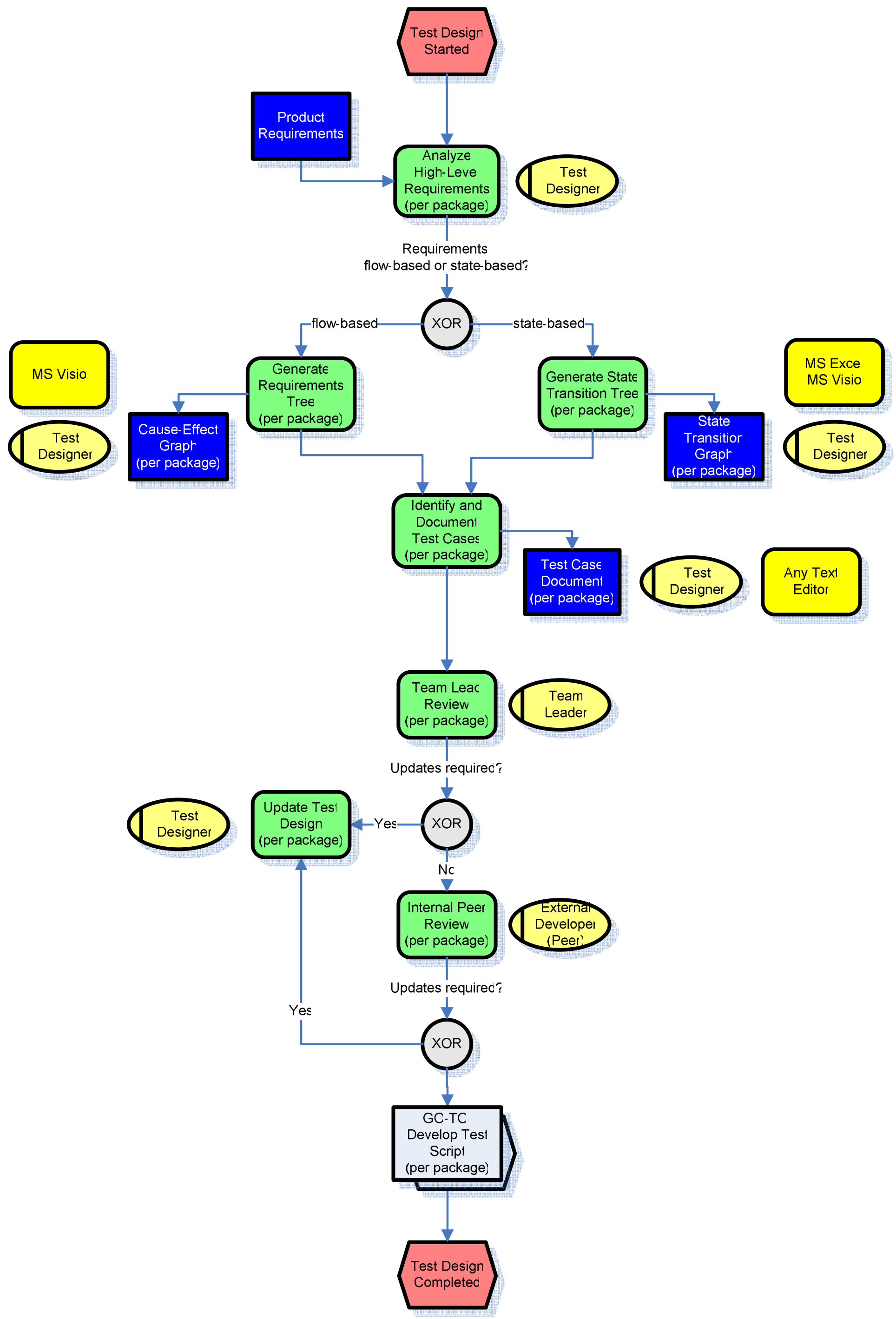


Figure 4.25 Flow of Test Design Process

The study was retrospective, and instead of identifying process attribute values to put on process similarity matrices by filling process execution records, we preferred drawing general process flows with a Team Leader (TL) this time. We depicted the executions as draft on a paper first together with the TL, and then converted the flows into MS Visio files using eEPC (Extended Event Driven Process Change) notation. The flow for test design process is given in Figure 4.25.

The elements (inputs, outputs, activities, roles, and tools) used to represent process flows showed us typical values of process attributes, which we put on the matrices and checked against completed process executions. The process similarity matrix for test design executions is provided in Figure 4.26. As seen from this figure, test design was not performed for the last 26 packages. Test script development was performed directly for these packages, all belonged to a separate module. We completed process similarity matrices for test script development and test development peer review processes as well.

			Process Executions																									
Process Attributes			PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	PE21	PE22	PE23	PE24	PE25	PE26
1	Inputs																											
	1.1	Product Requirements	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2	Outputs																											
	2.1	Requirements Tree										o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	2.2	Cause-Effect Graph										o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	2.3	State-Transition Graph	o	o	o	o	o	o	o	o	o																	
	2.4	Test Case Document	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3	Activities																											
	3.1	Analyze HL Requirements	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	3.2	Generate Test Design	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	3.3	Identify and Document Test Cases	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	3.4	Team Lead Review	o	o					o								o		o			o	o	o				
	3.5	Internal Peer Review	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	3.6	Update Test Design after IP Review	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4	Roles																											
	4.1	Test Designer	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	4.2	Team Leader	o	o					o								o		o			o	o	o				
	4.3	External Developer (Peer)	o	o	o		o	o	o	o							o		o			o	o	o				
5	Tools and Techniques																											
	5.1	MS Visio	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	5.2	MS Excel	o	o	o	o	o	o	o	o	o																	
	5.3	Text Editor	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Process Cluster																												

Figure 4.26 Process Similarity Matrix for Test Design

After finalizing the process similarity matrices, we analyzed them for similarity and differences in process executions. We identified 6 process clusters labeled from “a” through “f” as shown in Figure 4.27, by observing the similarities between process executions. The values of cluster distances were high, meaning that the clusters were not very similar to each other. The number of data points was few, and we decided to allow a cluster distance value of 3 for possible merges and identified 2 clusters accordingly: 1) a, b, c, d; and 2) e, f. When we looked at corresponding data, we noticed that these clusters represented two different modules under test design. Similarly, we investigated clusters for test script development and test development peer review processes but could not detect any, since the executions of these processes were very consistent. Therefore we used them as is. The list of processes and detected clusters are provided in Table 4.8.

			Process Executions					
Process Attributes			PE1	PE3	PE4	PE9	PE10	PE15
1	**Inputs**							
	1.1	Product Requirements	o	o	o	o	o	o
2	**Outputs**							
	2.1	Requirements Tree					o	o
	2.2	Cause-Effect Graph					o	o
	2.3	State-Transition Graph	o	o	o	o		
	2.4	Test Case Document	o	o	o	o	o	o
3	**Activities**							
	3.1	Analyze HL Requirements	o	o	o	o	o	o
	3.2	Generate Test Design	o	o	o	o	o	o
	3.3	Identify and Document Test Cases	o	o	o	o	o	o
	3.4	Team Lead Review	o					o
	3.5	Internal Peer Review	o	o		o	o	o
	3.6	Update Test Design after IP Review	o	o		o	o	o
4	**Roles**							
	4.1	Test Designer	o	o	o	o	o	o
	4.2	Team Leader	o					o
	4.3	External Developer (Peer)	o	o				o
5	**Tools and Techniques**							
	5.1	MS Visio	o	o	o	o	o	o
	5.2	MS Excel	o	o	o	o		
	5.3	Text Editor	o	o	o	o	o	o
Process Cluster			a	b	c	d	e	f

a-b	2								
a-c	5	b-c	3						
a-d	3	b-d	1	c-d	2				
a-e	7	b-e	5	c-e	6	d-e	4		
a-f	4	b-f	6	c-f	9	d-f	7	e-f	3

Cluster	Mergable To	Will Be Merged To	CLUSTER
a	-	-	1
b	a	a	1
c	b	b	1
d	a,b,c	b	1
e	-	-	2
f	e	e	2

Figure 4.27 Initial Process Clusters for Test Design Process

Table 4.8 Processes and Process Clusters Identified for Cases C, D, E (Context-3)

Case Study	Process	Cluster	Explanation
C	Test Design	1 2	Module #1 Module #2
D	Test Script Development	1	Original, of three modules
E	Test Development Peer Review	1	Original, of three modules

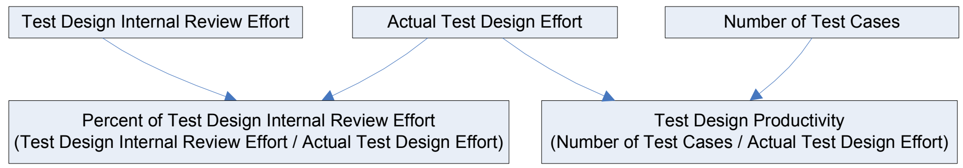
In addition to the clusters listed in the table above, we also derived data sets for test development and overall reviews. Test cases for Module-3 had been implemented directly by performing test script development and without performing test design. Since we wanted to evaluate overall performance of test development process among three modules, we added the effort of test design to that of test script development, and gathered total test development effort accordingly. On the other hand, to rationalize the number of action items detected in test development peer reviews, we wanted to evaluate total review effort spent for each package. We gathered overall review effort by adding internal review effort spent for test development to the peer review effort, for each package. The processes and data sets (both original and derived) subject to evaluation in our cases are listed in Table 4.9.

Table 4.9 Processes and Data Sets (Original and Derived) in Context-3

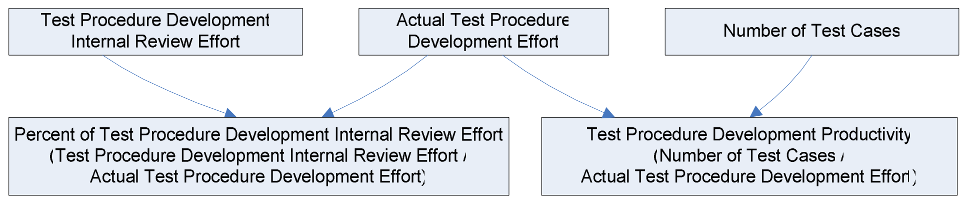
Process	**Data Source**	**Data (Collected or Gathered)**
Test Design	Original	Test design effort Test design internal review effort
Test Script Development	Original	Test script development effort Test script development internal review effort
Test Development	Derived	Test design effort + test script development effort
Test Development Peer Review	Original	Test development peer review effort Number of action items detected in the review Test development review update effort
Overall Reviews	Derived	Test design internal review effort + Test script development internal review effort + Test development peer review effort

After we identified initial process clusters, we worked on process metrics based on available data described in the table above, to evaluate usability for statistical analysis. The list of base and derived metrics we identified for each process as well as their formulas are given in Figure 4.28. The arrows show the relationships between the base metrics at upper side to the derived metrics at lower side.

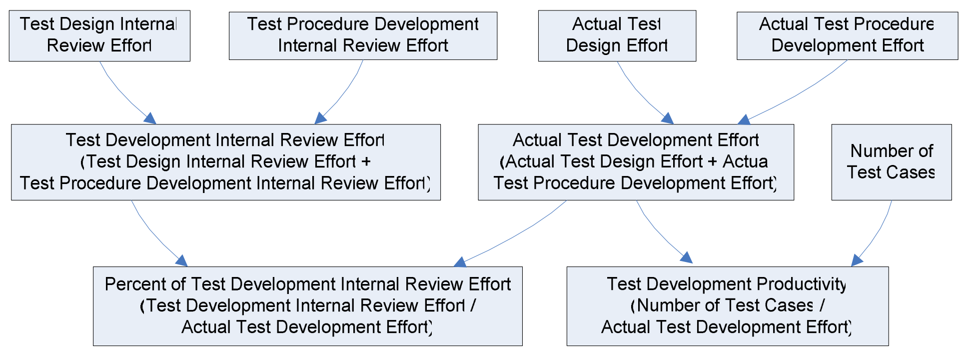
We filled Metric Usability Questionnaire for each base and derived metric shown in Figure 4.28. Example questionnaire for "Actual Test Design Effort" base metric is shown in Figure 4.29 (completed questionnaires for all metrics identified in Context-3 are provided in appendices E, F, and G). The usability status of all base and derived metrics are listed in Table 4.10. As seen from the table, all metrics were evaluated as "usable" and therefore would be used for control charting.



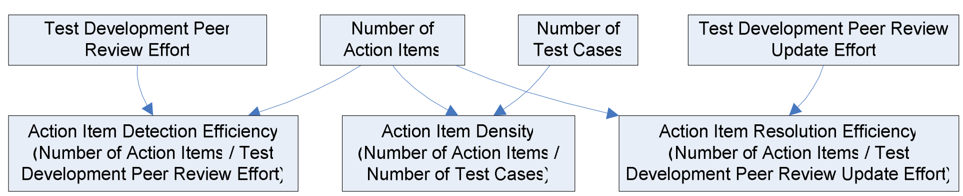
(a) Base and derived metrics identified for Test Design process



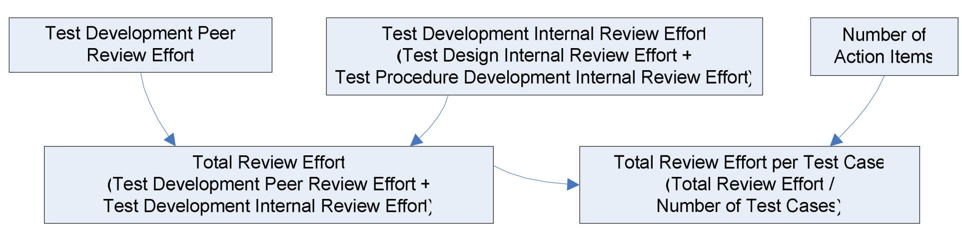
(b) Base and derived metrics identified for Test Procedure Development process



(c) Base and derived metrics identified for Test Development



(d) Base and derived metrics identified for Test Development Peer Review process



(e) Base and derived metrics identified for Overall Reviews

Figure 4.28 Base and Derived Metrics Identified in Context-3

Please rate each attribute in four scales, based on asnwers to questions as indicators:
F : Indicatiors of the atribute are fully satisfied (%86-100)
L : Indicatiors of the atribute are largely satisfied (%51-85)
P : Indicatiors of the atribute are largely satisfied (%16-50)
N : Indicatiors of the atribute are not satisfied (%0-15)

Metric Name: Actual Test Design Effort
Conceptual Definition: Total actual effort spent for test design of a package
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Attributes		Answers		Rating	Expected Answers
Indicators					
Metric Identity			**MUF-1**	**F**	
Q1	Which entity does the metric measure?	Human Resource			
Q2	Which attribute of the entity does the metric measure?	Effort			
Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio			Ratio, Absolute
Q4	What is the unit of the metric data?	Person-hours			
Q5	What is the type of the metric data? (integer, real, etc.)	Real			
Q6	What is the range of the metric data?	[6.00, 223.00]			
Data Existence				**F**	
Q7	Is metric data existent?	Yes			Available > 20
Q8	What is the amount of overall observations?	26			
Q9	What is the amount of missing data points?	1			
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No			
Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.			
Data Verifiability			**MUF-3**	**F**	
Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end			
Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes			Yes
Q14	Who is responsible for recording metric data?	Test Designer			
Q15	Is all metric data recorded by the responsible body?	Yes			Yes
Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet			
Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes			Yes
Q18	Where is metric data stored? (in a file, database, etc.)	In a file			
Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes			Yes
Data Dependability				**L**	
Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously			
Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously			
Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously			
Q23	Are the frequencies for data generation, recording, and storing different?	No			No
Q24	Is metric data recorded precisely?	Yes			Yes
Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)			Yes
Q26	Is the purpose of metric data collection known by process performers?	Yes			Yes
Q27	Is metric data analyzed and reported?	No			Yes
Q28	Is metric data analysis results communicated to process performers?	No			Yes
Q29	Is metric data analysis results communicated to management?	No			Yes
Q30	Is metric data analysis results used as a basis for decision making?	No			Yes
Data Normalizability					
Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)			
Data Integrability					
Q32	Is metric data integrable at project level?	Yes			
Q33	Is metric data integrable at organization level?	No			

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure 4.29 Metric Usability Questionnaire for “Actual Test Design Effort” Base Metric of “Test Design” Process

We reviewed process data and used the results from process similarity assessment and metric usability evaluation to finalize process clusters and metrics prior to control charting. After the review, we decided to identify a cluster for each of the three modules. This decision was conformant with the result of process similarity assessment performed for Test Design shown in Table 4.8. For other processes, we sub-clustered the original executions with respect to the modules (as M1, M2, and M3), to enable comparison among them within the project. In other words, we wanted to understand variations, if any, in performances of test design, test script development, and test design peer review processes among the modules of the project. We should note that test development and review for Module-1 and Module-2 were performed by the same group while test development and review for Module-3 were performed by another group. We should also note that the second group did not perform test design and directly developed test scripts. We thought the results would provide information on the performances of these two groups as well.

Table 4.10 Metric Usability Results in Context-3

Metric	**Type**	**Usability Status**
Actual Test Design Effort	Base	U (Largely)
Test Design Internal Review Effort	Base	U (Largely)
Number of Test Cases	Base	U (Largely)
Test Design Productivity	Derived	U (Largely)
Percent of Test Design Internal Review Effort	Derived	U (Largely)
Actual Test Procedure Development Effort	Base	U (Largely)
Test Procedure Development Internal Review Effort	Base	U (Largely)
Test Procedure Development Productivity	Derived	U (Largely)
Percent of Test Procedure Development Internal Review Effort	Derived	U (Largely)
Test Development Peer Review Effort	Derived	U (Largely)
Test Development Peer Review Update Effort	Derived	U (Largely)
Test Development Peer Review Effort	Base	U (Largely)
Test Development Peer Review Update Effort	Base	U (Largely)
Number of Action Items	Base	U (Largely)
Action Item Detection Efficiency	Derived	U (Largely)
Action Item Resolution Efficiency	Derived	U (Largely)
Action Item Density	Derived	U (Largely)
Total Review Effort	Derived	U (Largely)
Total Review Effort per Test Case	Derived	U (Largely)

Table 4.11 Derived Metrics Utilized to Understand Process Performances in Context-3

Process	**Derived Metric**	**Proc.Clusters**
Test Design	Test Design Productivity	M1, M2
	Percent of Test Design Internal Review Effort	M1, M2
Test Procedure Development	Test Procedure Development Productivity	M1, M2, M3
	Percent of Test Procedure Development Internal Review Effort	M1, M2, M3
Test Development	Test Development Productivity	M1, M2, M3
	Percent of Test Development Internal Review Effort	M1, M2, M3
Test Development Peer Review	Action Item Density	M1, M2, M3
	Action Item Detection Efficiency	M1, M2, M3
	Action Item Resolution Efficiency	M1, M2, M3
Overall Reviews	Total Review Effort per Test Case	M1, M2, M3

We first charted combined data for each process, and then charted the data of each cluster (module) separately. Table 4.11 shows derived metrics utilized to understand the performance of each process. We applied variables charts for individuals of metrics data.

As an example, the chart for combined data of test design productivity is given in Figure 4.30. The charts for clusters M1 and M2 for the same metric are shown in Figure 4.31. Control charts for all derived metrics and the clusters listed in Table 4.11 are provided in the appendices E, F, and G.

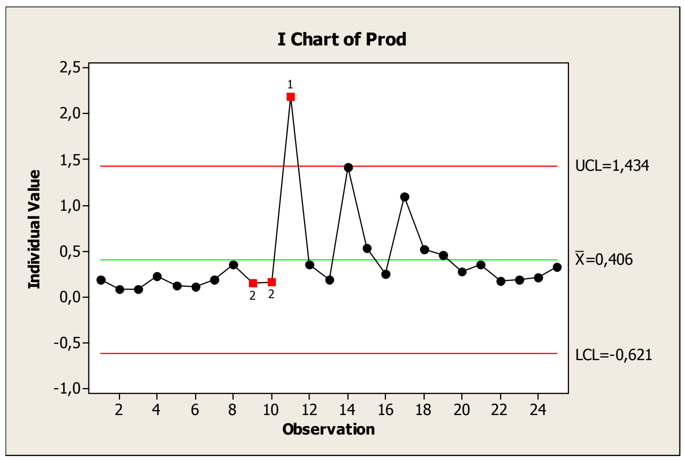
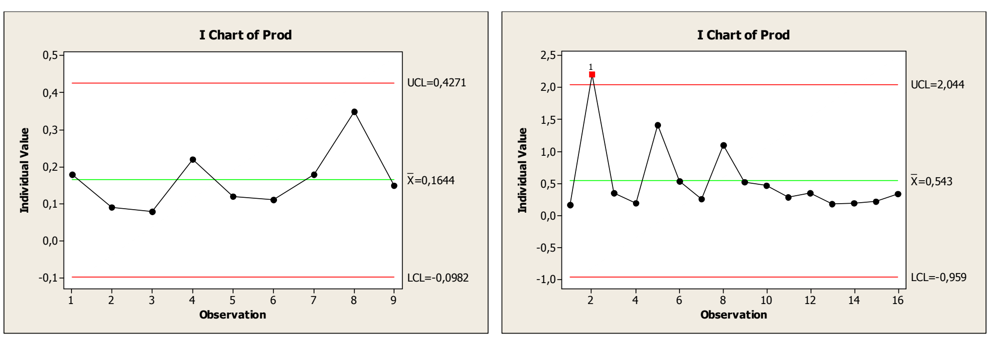


Figure 4.30 Individuals Chart for Combined Data of Test Design Productivity



(a) Test design productivity of cluster M1 (b) Test design productivity of cluster M2

Figure 4.31 Individuals Charts for Test Design Productivity

From the figure above we saw that cluster M1 was under control, and cluster M2 had an out-of-control point with respect to test design productivity metric. Initial results obtained from all control charts are summarized in Table 4.12.

Table 4.12 Initial Results from Charted Data in Context-3

Process	**Metric**	**Cluster**	**Status**
Test Design	Test Design Productivity	Overall	3 OCPs
		M1	Under Control
		M2	1 OCP
	Percent of Test Design Internal Review Effort	Overall	2 OCPs
		M1	Under Control
		M2	Under Control
Test Procedure Development	Test Procedure Development Productivity	Overall	Many OCPs
		M1	Under Control
		M2	1 OCP
		M3	1 OCP
	Percent of Test Procedure Development Internal Review Effort	Overall	5 OCPs
		M1	1 OCP
		M2	Under Control
		M3	Under Control
Test Development	Test Development Productivity	Overall	4 OCPs
		M1	Under Control
		M2	1 OCP
		M3	1 OCP
	Percent of Test Development Internal Review Effort	Overall	Under Control
		M1	Under Control
		M2	Under Control
		M3	Under Control
Test Development Peer Review	Action Item Density	Overall	5 OCPs
		M1	Under Control
		M2	Under Control
		M3	1 OCP
	Action Item Detection Efficiency	Overall	4 OCPs
		M1	Under Control
		M2	Under Control
		M3	Under Control
	Action Item Resolution Efficiency	Overall	Many OCPs
		M1	Under Control
		M2	Under Control
		M3	Under Control
Overall Reviews	Total Review Effort per Test Case	Overall	Many OCPs
		M1	Under Control
		M2	2 OCPs
		M3	Under Control

* OCP: Out-of-Control Point

We conducted an interview with the Team Leader in order to understand any reasons for the assignable causes. The reasons were investigated by filling process execution questionnaire for each out-of-control point reported by the table above. Process execution questionnaire for process execution-1 is shown in Figure 4.32.

Process Name: Test Script Development		Recorded On: 21.05.2006	
Process Execution No: 1		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	No	The first package under development
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The package has passed many revisions, it was much like an example package. It enabled on-the-job training.		

Figure 4.32 Process Execution Questionnaire for Test Procedure Development Process Execution # 1

Detailed findings of the reasons for out-of-control points as detected by process execution questionnaires are given in Table 4.13. During the interview, we further asked for likely causes behind out-of-control points for which no specific reason could be detected by the questionnaires. We could only find an answer for "total review effort per test case" metric for Module-2. When we looked at the data set with the Team Leader, he pointed a specific package having a high complexity and 9 distinct layers. The layers had similar parts so that "copy-paste" was utilized during test development, leading repetition of defects. Therefore review time increased to find out these defects. We excluded related data point from the set (regarding PE23) and re-charted remaining data; and observed that the cluster came under control.

Based on the knowledge obtained during the interview, we re-charted the data by excluding the out-of-control points, all having assignable causes. Final results are summarized in Table 4.14.

Table 4.13 Assignable Causes for Out-of-Control Points in Context-3

Metric	Cluster	OCPs	Assignable Cause
Test Design Productivity	M-1	None	Not applicable
	M-2	1	High productivity due to a very experienced test designer who already had domain knowledge
Percent of Test Design Internal Review Effort	M-1	None	Not applicable
	M-2	None	Not applicable
Test Procedure Development Productivity	M-1	None	Not applicable
	M-2	1	High productivity due to a very experienced test developer who already had domain knowledge
	M-3	1	No specific reason detected by PEQ
Percent of Test Procedure Development Internal Review Effort	M-1	1	High internal review percent due to being the first package under development
	M-2	None	Not applicable
	M-3	None	Not applicable
Test Development Productivity	M-1	None	Not applicable
	M-2	1	High productivity due to a very experienced test developer who already had domain knowledge
	M-3	1	No specific reason detected by PEQ
Percent of Test Development Internal Review Effort	M-1	None	Not applicable
	M-2	None	Not applicable
	M-3	None	Not applicable
Action Item Density	M-1	None	Not applicable
	M-2	None	Not applicable
	M-3	1	High defectiveness because the package under review was developed by a very inexperienced staff
Action Item Detection Efficiency	M-1	None	Not applicable
	M-2	None	Not applicable
	M-3	None	Not applicable
Action Item Resolution Efficiency	M-1	None	Not applicable
	M-2	None	Not applicable
	M-3	None	Not applicable
Total Review Effort per Test Case	M-1	None	Not applicable
	M-2	2	No specific reason detected by PEQs
	M-3	None	Not applicable

* OCP: Out-of-Control Point, PEQ: Process Execution Questionnaire

Table 4.14 Final Results from Charted Data in Context-3

	Process Clusters		
Metric	**M1**	**M2**	**M3**
Test Design Productivity	Under Control	1 OCPs	-
Percent of Test Design Internal Review Effort	Under Control	Under Control	-
Test Procedure Development Productivity	Under Control	Under Control	1 OCP
Percent of Test Procedure Development Internal Review Effort	Under Control	Under Control	Under Control
Test Development Productivity	Under Control	1 OCP	1 OCP
Percent of Test Development Internal Review Effort	Under Control	Under Control	Under Control
Action Item Density	Under Control	Under Control	Under Control
Action Item Detection Efficiency	Under Control	Under Control	Under Control
Action Item Resolution Efficiency	Under Control	Under Control	Under Control
Total Review Effort per Test Case	Under Control	Under Control	Under Control

From the table above we observed that M1 was under control with respect to all metrics of all processes. M2 had an out of control point without an assignable cause with respect to test design productivity. Similarly, M3 had an out of control point without an assignable cause with respect to test procedure development productivity. From the process point of view, on the other hand, we observed that all review metrics (regarding internal reviews, peer reviews and overall reviews) were under control. However, productivity metrics (regarding test design, test procedure development, and test development) were not under control.

Findings from the study:

After detecting the control status of process clusters with respect to specified metrics, we compared mean values of metric data to quantify our findings. Mean values of process metrics for each cluster are given in Table 4.15.

Table 4.15 Mean Values of Process Metrics for Each Cluster in Context-3

		Process Cluster Means		
Process	**Metric**	**M1**	**M2**	**M3**
Test Design	Test Design Productivity	0,16	0,43	-
	Percent of Test Design Internal Review Effort	0,09	0,28	-
Test Procedure Development	Test Procedure Development Productivity	0,15	0,18	0,07
	Percent of Test Procedure Development Internal Review Effort	0,07	0,08	0,18
Test Development	Test Development Productivity	0,07	0,11	0,07
	Percent of Test Development Internal Review Effort	0,09	0,15	0,18
Test Development Peer Review	Action Item Density	0,74	2,66	5,14
	Action Item Detection Efficiency	1,14	2,84	1,09
	Action Item Resolution Efficiency	1,71	38,7	1,39
Overall Reviews	Total Review Effort per Test Case	1,95	2,64	11,98

Based on the data on the table above, we had the following findings among process clusters:

- Test design productivity for M2 was nearly three times that of M1. However, percent of test design internal review effort and action item density for M2 were three times those of M1. This was a trade-off. Productivity was high but the cost of quality was also high for M2. (We should note that the cost of quality is measured by the cost of achieving quality added to cost of not achieving quality. Here we accepted the percent of internal review effort as the cost of achieving quality and action item density as the cost of not achieving quality while having judgment on the cost of quality.)
- Test procedure development productivity was the lowest for M3 but percent of test procedure development internal review effort was the highest. We attributed this to not performing test design for M3.
- Test development productivity and the percent of test development internal review effort for M1 had the lowest values among three clusters. Action item density for M1 was also the lowest. M1 was in steady state. We should remind that M1 was under control for all metrics of all processes.

- Test development productivity for M3 was the same as that of M1, and percent of test development internal review effort for M3 was twice that of M1. Test design (and its internal reviews) was performed for M1 but not for M3. It was interesting that percent of internal review effort for M3 was higher than that of M1. When we looked at action item density we could resolve why this is the case. Action item density for M3 was seven times that of M1. It was obvious that performing test design for M1 enabled early detection of defects, decreasing both action item density of peer review and percent of test development internal review effort of test development.
- Action item density of peer review for M3 was the highest. Also, M3 had the lowest productivity and the highest percent for test development internal review effort. Therefore, M3 was the worst package in terms of productivity and quality. If we consider percent of internal review effort and action item density, we can say that M3 has the highest cost of quality for the project.
- Action item detection efficiency of peer review for M2 was twice that of others, and action item resolution efficiency of peer review for M2 was extremely high. We attributed this to the experience and domain knowledge of the staff worked in test development of the cluster.
- Total review effort per test case for M3 was very high (10 times that of M1 and 5 times that of M2). This was despite the fact that test design (and its internal reviews) was not performed for M3. It was obvious that internal reviews during test procedure development and the peer review demanded more effort for the cluster.

CHAPTER 5

CONCLUSION

Statistical process control offers tools for controlling software processes to manage projects with allowed variation, to refine product quality, or to improve process capability. This valuable set of tools is encouraged in the industry mostly by adopting organizational maturity or process capability frameworks that are offered by process improvement models. These models demand investment of time, effort, and money for several years, which are difficult to afford for emergent organizations that need more practical methods to understand and manage their processes based on quantitative data. To address this difficulty, we developed an assessment model to test the suitability of SPC for software processes and metrics.

The model describes procedure and assets for understanding the context in which process data are generated, identifying rational samples of process executions based on this understanding, identifying metrics to use for statistical analysis according to metric and data characteristics, and determining the level of confidence of success in SPC implementations via control charts. While identifying rational samples, process performers have a broad understanding of process and its attribute values at process executions. This understanding is crucial in interpreting data analysis results and can be attained without the need to explicitly define the process. While identifying metrics to use for statistical analysis, process performers discover the characteristics of metric data and investigate performance of basic measurement practices during collection of data. The performance of measurement practices provides information on the health of data collected, and in turn, of data analysis results. This information is attained by answering a questionnaire for the metric under study and without the need for defining an explicit measurement process. Such characteristics of the assessment model provide practicality and flexibility especially for emergent organizations that may not have an established process management infrastructure but are willing to understand the performance of their practices based on quantitative data. The existence of process definitions, including the measurement process definition, is welcomed and can make the assessment process easier, although not essential.

To refine and validate our model, we performed three case studies in a multiple-case-study context. For each of the case studies, we identified rational samples of the process under study via process similarity assessment, and evaluated the usability of candidate process metrics by using questionnaires and a rating scheme. We worked on task management, review, test development processes and related metrics of three organizations. The first case investigated utilization of estimation capability and effort variance metrics of task management process of a project-based working software organization. In the second case, we worked on non-conformance detection efficiency, non-conformance resolution efficiency, review open period, and review open period with respect to non-conformances metrics of review process of a system and software development organization targeting to achieve CMMI L3. In the third case, we worked on test design, test procedure development, and test development peer review processes of a system and software development organization having SW-CMM L3. We investigated the utilizations of productivity and percent of internal review effort metrics for test design and test procedure development processes, and the utilizations of action item density, action item detection efficiency, and action item resolution efficiency metrics for test development peer review process.

For task management process (case study-A), our model suggested that both estimation capability and effort variance derived metrics could be used for statistical analysis; and two process clusters identified by process similarity assessment were detected as under control with respect to both derived metrics

("Under control" means all out-of-control points have been removed and prevented from recurring in the future). The ranges of estimation capability were wide for the process cluster in which verification activity had not been performed and we could not be sure on the validity of actual effort spent for these tasks. Accordingly, we recommended to the team leader to check open status of tasks regularly and to perform their closures on time as possible. We also suggested that task assignments should be updated as the overall project plan changes to maintain consistency between the work assignments. The ranges for effort variance were wide for process clusters and we attributed high variance to different sizes of work performed by the tasks. On the other hand, the control chart for effort variance of the process cluster in which verification activity had been performed showed that the difference between estimated and actual effort values decreased in time, meaning an improvement in the tasks management process. After case study-A, the project's team leader included reviews of task management data in regular progress monitoring in order to perform task colures on time and to update task assignments in consistency with the project plan.

For review process (case study-B), we identified two process clusters and evaluated that nonconformance detection efficiency, nonconformance resolution efficiency, and open period with respect to nonconformances metrics were largely usable for SPC analysis while open period metric was poorly usable. We decomposed one of the process clusters into four further sub-clusters by categorizing process data with respect to product type under review. After charting the data, we observed that all process clusters were under control with respect to nonconformance detection efficiency and nonconformance resolution efficiency metrics. Charted data showed that the company could not use the control charts for open period with respect to nonconformances metric confidently, although the metric was evaluated as usable by our approach. After the interviews that we performed to detect the assignable causes, we found out that the schedule of the projects played a significant role in the open periods of review records. To evaluate review process performance, we utilized existing LOC data to rationalize nonconformance detection efficiency. We calculated nonconformance density metric by the formula "number of accepted nonconformances/KLOC" and we charted the metric data: Both process clusters were under control with respect to nonconformance density metric, meaning that nonconformance detection efficiency could be used to judge and improve process performance. We noted that the company should keep recording product size to continually monitor nonconformance density for possible changes in the performance, and recommended recording the size of the product under review on process assets in each review. On the other hand, nonconformance resolution efficiency metric could be useful for process performers to estimate the effort required for resolving the nonconformances, if the product was to be updated after review meeting. The type of nonconformance could have a significant effect on estimating nonconformance resolution effort since syntactic errors would take less effort to fix while semantic ones would take more. Accordingly, we recommended recording the type of each nonconformance accepted in each review. After case study-B, process performers started to record product size on review records. The measurement representative initiated SPC implementations for nonconformance detection efficiency and nonconformance resolution efficiency metrics, and adopted related control charts as part of the measurement and analysis system built for CMMI Level 3. By doing so, the company had the chance of observing and improving review process performance based on quantitative data, which is a basic requirement for achieving higher CMMI maturity levels.

For test design (case study-C), test procedure development (case study-D), and test development peer review (case study-E), processes; we identified a process cluster for each of the three modules and all process metrics were evaluated as "largely usable" for statistical analysis. After charting the data and removing the assignable causes, we observed that; the first module was under control with respect to all metrics of all processes, the second module had an out of control point without an assignable cause with respect to test design productivity, and the third module had an out of control point without an assignable cause with respect to test procedure development productivity. From the process point of view, on the other hand, we observed that; all review metrics (regarding internal reviews, peer reviews and overall reviews) were under control, but productivity metrics (regarding test design, test procedure development, and test development) were not under control. Therefore the suggestions of our model were confirmed except for the productivity metrics. Accordingly, we recommended to the Team Leaders that productivity of test design, test procedure development and test development should be monitored in close relation with the review metrics, or in other words, with the cost of quality of resulting test packages. After the case studies, we had a meeting with the process improvement team leader and measurement representative and decided to share our findings with project manager and project team in order to receive their feedback and rationales. We expect the findings can be used for institutionalization of test design and test development processes as well as project (re)planning.

We should note that we almost always used the assets of the model easily. Only for the first case, we had some difficulty in identifying process clusters and evaluating metric usability due to lack of established rules. However, we resolved this difficulty for the following cases by defining the rules both for identification of process clusters and for evaluation of metric usability.

SPC trials helped to set and refine the understanding of the issues under study (project/process performance, product quality, etc.) in all case studies. We clearly observed that the acts of measuring and analyzing are themselves a means of process improvement. While trying to chart data and interpret chart results, we (together with process performers) checked and refined our understanding of process executions and their contextual information. Trying to identify and eliminate the assignable causes enabled detailed study of individual process executions. Refining process understanding naturally brought recommendations for improvement. At the end of the case studies, the organizations updated their process definitions and assets based on our findings. For example in case study-A, the team leader included reviews of task management data in regular progress monitoring; and in case study-B, review process owner updated the review record to keep the size of each product under review.

We spent 46 man-hours for case study-A, 115 man-hours for case study-B, and 44 man-hours for case study-C. The effort values are considerably small when compared with the effort of process performances (e.g., %6 for case study-B).

Accordingly we may say the application of the assessment model does not require a high effort.

During the case studies, metric usability evaluation enabled us to select metrics that will succeed in statistical analysis (not only with control charts). This is especially important for software organizations that are unfamiliar to but feel strong need to use statistical techniques. Though, we observed that evaluating usability of metrics was supporting but not enough to effectively select the metrics to be used in SPC analysis. Project context and dynamics in which the process was executed (such as project organization, schedule, development life cycle, maturity of development practices, and etc.) should also be considered while selecting the metrics. For example, re-charted data showed in case study B that the company could not use the control charts for open period with respect to nonconformances metric confidently, although the metric was evaluated as usable by our approach. After the interviews we could detect that the schedule of the projects played a significant role in the open periods of review records. Elaboration on process metrics prior to SPC implementation requires special attention from this perspective. We can work on each process metric specifically and investigate factors that might affect its utilization.

By process consistency assessment we could systematically identify rational samples of the process, which is difficult to achieve especially in software engineering domain. This is very important to satisfy the basic requirement of achieving process control: “Build single and constant system of chance causes”. We observed that identification of rational samples is closely related to the purpose of SPC implementation (though we did not specifically consider the purpose of implementation in the studies since we were primarily trying to validate our model). In case study A, we tried to understand the effects of task types in task management of a project. In case study B, we worked to identify different executions of a review process organization-wide. For case studies C, D, and E, we again worked in project-context, and compared test development performance and quality for three different modules. Therefore, if we had defined the purpose for all these cases at the beginning, we would have used different phrases for each, and identified process clusters accordingly. Selection of the metrics would also have been affected by these phrases.

We have a number of constraints related to the assessment model and its applications. The first one is retrospective characteristic of the case studies. We questioned the attributes of past executions and since we worked on existing process executions and assets, we had difficulties in catching implementation details. Organizing a prospective case study will support better understanding of process executions and related characteristics. Second, we performed the assessments by ourselves and by consulting process performers, and we could not verify whether the model is easily usable by the company staff. Especially the use of metric usability questionnaire to judge a metric's usability has subjectivity in some parts and requires expertise in measurement and analysis concepts. Accordingly, developing more specific guidelines that describe how to perform process consistency assessment and metric usability evaluation might be useful. The third constraint is that metric usability evaluation provides information on a metric's usability for statistical analysis, but it does not state whether the utilization would be effective. The selection of effective metrics for a process needs further elaboration. Fourth, we cannot generalize the results from our case studies since the variety in the type of organizations is limited. The organization of case study-B had established processes and has been pursuing studies towards CMMI Level 3 certification, and the organization of case study-C had already CMM Level 3. The usability of the model needs to be tested by conducting more case studies in various contexts. These constraints also show the directions for the future work.

As a result, multiple case study implementations showed us that our model utilized for rational sampling and metric selection was useful as a guide for starting SPC implementation in emergent organizations. The first question we were investigating was "Can we identify approaches to direct SPC implementation?" and we defined an assessment model for this purpose. The second question was "Can an emergent organization apply SPC techniques following these approaches and benefit from the results?" and we applied the assessment model in three emergent organizations that all benefited from the applications. The only deficiency was the applying body: “us” instead of the “company staff”.

We are aware that starting SPC implementation is not enough and success demands for continuous monitoring and cause analysis to improve process capability. Our model has served as a good vehicle to set the ground for such efforts. We expect software companies quickly adopt SPC techniques by using The Assessment Model for Statistical Process Control.

REFERENCES

[1] Fenton, N.E., and Pfleeger, S.L., Software Metrics: A Rigorous and Practical Approach (2nd Ed.). PWS Publishing Company, 1997.

[2] CMU/SEI, Capability Maturity Model Integration – Version 1.1. Technical Report (Continuous: CMU/SEI-2002-TR-001, Staged: CMU/SEI-2002-TR-002), December 2001.

[3] ISO/IEC JTC1/SC7, ISO/IEC TR 15504: 1998: Information Technology – Software Process Assessment.

[4] Montgomery, D.C., Introduction to Statistical Quality Control (5^{th} Ed.). John Wiley & Sons, Inc., U.S.A., 2005. ISBN 0-471-65631-3.

[5] Shewhart, W.A., Economic Control of Quality of Manufactured Product. Van Nostrand, New York, 1931. (Re-printed by American Society of Quality Control, Milwaukee, Wisc., 1980.)

[6] Deming, W.E., Statistical Adjustment of Data. John Wiley and Sons, 1943. (Re-printed by Dover Publications, July 1984.)

[7] Deming, W.E., Out of the Crisis. Massachusetts Institute of Technology, Center of Advanced Engineering, Cambridge, Mass., 1986.

[8] Lantzy, M.A., "Application of Statistical Process Control to Software Processes". Proceedings of the Ninth Washington Ada Symposium on Empowering Software Users and Developers, 1992, pp.113-123.

[9] Card, D., “Statistical Process Control for Software?”. IEEE Software, pp.95-97, May 1994.

[10] Barnard, J., and Carleton, A., “Analyzing a Mature Software Inspection Process Using Statistical Process Control” (presentation). National SEPG Conference, Pittsburgh, 1999.

[11] Florac, A.W., Carleton A.D., and Barnard, J.R., "Statistical Process Control: Analyzing a Space Shuttle Onboard Software Process". IEEE Software, July/August 2000, pp.97-106.

[12] Weller, E.F., "Practical Applications of Statistical Process Control". IEEE Software, May/June 2000, pp.48-55.

[13] Jalote, P., Dinesh, K., Raghavan, S., Bhashyam, M.R., and Ramakrishnan, M., "Quantitative Quality Management through Defect Prediction and Statistical Process Control". Proceedings of Second World Quality Congress for Software, September 2000.

[14] Jacob, L., and Pillai, S.K., "Statistical Process Control to Improve Coding and Code Review". IEEE Software, May/June 2003, pp.50-55.

[15] Demirörs, O., and Sargut, K.U., "Utilization of a Defect Density Metric for SPC Analysis". Proceedings of the 13th International Conference on Software Quality, Dallas, Texas, October 2003.

[16] Demirörs, O., and Sargut, K.U., "Utilization of Statistical Process Control (SPC) in Emergent Software Organizations: Pitfalls and Suggestions". Software Quality Journal, Vol.14, No.2, pp.135-157, 2006.

[17] Carleton, A.D., and Paulk, M.C., "Can Statistical Process Control Be Usefully Applied To Software?" (presentation). The 4th European Software Engineering Process Group Conference, Amsterdam, June 1999.

[18] Florac, A.W., Carleton A.D., "Statistically Controlling the Software Process" (presentation). The 99 SEI Software Engineering Symposium, Software Engineering Institute, Carnegie Mellon University, September 1999.

[19] Florac, A.W., Park, R.E., and Carleton A.D., Practical Software Measurement: Measuring for Process Management and Improvement. Guidebook: CMU/SEI-97-HB-003, 1997.

[20] Paulk, M.C., "Considering Statistical Process Control for Software" (presentation). Software Engineering Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, September 2001.

[21] Radice, R., "Statistical Process Control for Software Projects". The 10th Software Engineering Process Group Conference, Chicago, Illinois, March 1998.

[22] Paulk, M.C., Weber, C.V., Curtis, B., and Chrissis, M.B., The Capability Maturity Model: Guidelines for Improving Software Process. Addison-Wesley Publishing, October 1995.

[23] CMU/SEI, "Process Maturity Profile of the Software Community – 2000 Year End Update" (presentation). March 2001.

[24] CMU/SEI, The 2001 High Maturity Workshop. CMU/SEI-2001-SR-014, January 2002.

[25] Grady, R.B., and Caswell, D.L., Software Metrics: Establishing a Company-Wide Program. Prentice Hall PTR, 1998.

[26] Hall, T., and Fenton, N., "Implementing Effective Software Metrics Programs". IEEE Software, 1997.

[27] Rozum, J.A., The SEI and NAWC: Working Together to Establish a Software Measurement Program. Technical Report, CMU/SEI-93-TR-07, ESC-TR-93-184, December 1993.

[28] Basili, V.R., "Software Modeling and Measurement: The Goal Question Metric Paradigm". Computer Science Technical Report Series, CS-TR-2956 (UMIACS-TR-92-96), University of Maryland, College Park, MD, September 1992.

[29] Basili, V.R., Caldiera, G., and Rombach, H.D., "The Goal Question Metric Approach". Encyclopedia of Software Engineering, Vol.1, pp.528-532, John Wiley & Sons, 1994.

[30] Park, R.E., Goethert, W.B., and Florac, W.A., Goal-Driven Software Measurement. CMU/SEI-96-HB-002, August 1996.

[31] Offen, R.J., and Jeffery, R., "Establishing Software Measurement Programs". IEEE Software, 1997.

[32] Argyris, C., "Good Communication that Blocks Learning". Harward Business Review, July-August 1994.

[33] Drucker, P.F., "The New Society Organizations". Harward Business Review, September-October, 1992.

[34] Hetzel, W.C., Making Software Measurement Work: Building an Effective Software Measurement Program. QED Publishing, Wellesley, Mass., 1993.

[35] Florac, A.W., Carleton A.D., Measuring the Software Process: Statistical Process Control for Software Process Improvement. Pearson Education, 1999 (a). ISBN 0-201-60444-2.

[36] Ishikawa, K., Guide to Quality Control. Asian Productivity Organization, 1982. ISBN 92-833-1035-7.

[37] Wheeler, D.J., Advanced Topics in Statistical Process Control. SPC Press, Knoxville, Tenn., 1995.

[38] Burr, A., and Owen, M., Statistical Methods for Software Quality. International Thomson Computer Press, 1996.

[39] Sargut, U., Application of Statistical Process Control to Software Development Processes. Master Thesis Dissertation, Informatics Institute of METU, May 2003.

[40] Florence, A., "Statistical Process Control Applied to Software Requirements Specification Process" (presentation). The 10th European Software Engineering Process Group Conference, London, June 2005.

[41] McGarry, J., Card, D., Jones, C., Layman, B., Clark, E., Dean, J., and Hall, F., Practical Software Measurement: Objective Information for Decision Makers (1st Ed.). Addison-Wesley Professional, 2001. ISBN 0201715163.

[42] ISO/IEC JTC1/SC7, ISO/IEC 15939: 2002: Software Engineering – Software Measurement Process.

[43] Paulk, M.C., "Practices for High Maturity Organizations". Proceedings of the 1999 Software Engineering Process Group Conference, Atlanta, Georgia, March 1999, pp.28-31.

[44] Yin, R.K., Case Study Research: Design and Methods. Applied Social Research Methods Series Vol.5, SAGE Publications, 2003.

[45] ISO, ISO 9001: 2000: Quality Management Systems – Requirements.

[46] NATO, AQAP-150: 1997: NATO Quality Assurance Requirements for Software Development (Edition 2).

[47] MINITAB Statistical Software (Release 14). http://www.minitab.com/products/minitab/14/ default.aspx.

APPENDICES

APPENDIX A. ASSESSMENT ASSETS

We developed a number of assets to use while performing assessments in accordance to SPC-AM. The list of assessment assets are given below. The assets themselves are shown in the following pages.

- Process Execution Record
- Process Execution Questionnaire
- Process Similarity Matrix
- Process Attributes Description
- Metric Usability Questionnaire for Base Metrics
- Metric Usability Questionnaire for Derived Metrics

SPC[AM]
Process Execution Record
(Internal Attributes)

Process Name:		Recorded On:	
Process Execution No:		Recorded By:	

1. **Inputs**: Please list the inputs to the process execution.

No	Name	Description

2. **Outputs**: Please list the outputs from the process execution.

No	Name	Description

3. **Activities (in sequence)**: Please list in sequence the activities that were performed while executing the process.

No	Name	Description

4. **Roles**: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description

5. **Tools and Techniques**: Please list the tools and techniques that are used to support process execution.

No	Name	Description

Figure A.1 Process Execution Record

Process Name:		Recorded On:	
Process Execution No:		Recorded By:	
External Attributes		Status (Yes/No)	Explanation
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?		
Q2	Are process performers experienced in their roles in the process?		
Q3	Are process performers differed per role basis during execution of the process?		
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?		
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)		
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)		
Q7	Has there been a recent change in funding and resources allocated for the process?		
Q8	Has the process been tailored for this specific execution?		
OTHER FACTORS (Please list if any)			

Figure A.2 Process Execution Questionnaire

PROCESS SIMILARITY MATRIX																				
Process Name:																				
Recorded On:																				
Recorded By:																				
		Process Executions																		
Process Attributes		PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	.	.	.	.	.	.	.	.	PEX
1	**Inputs**																			
	1.1																			
2	**Outputs**																			
	2.1																			
3	**Activities**																			
	3.1																			
	Activities in this sequence?																			
4	**Roles**																			
	4.1																			
5	**Tools and Techniques**																			
	5.1																			
Process Cluster																				
* Please verify each process execution against process attributes. Insert an "O" into each cell if applicable (leave blank if not applicable).																				

Figure A.3 Process Similarity Matrix

SPC

Process Attributes Description

Process Name:		Described On:	
Process Cluster:		Described By:	

1. **Inputs**: Please list the inputs to the process.

No	Name	Description

2. **Outputs**: Please list the outputs from the process.

No	Name	Description

3. **Activities (in sequence)**: Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description

4. **Roles**: Please list the roles that are allocated responsibilities in the process.

No	Name	Description

5. **Tools and Techniques**: Please list the tools and techniques that are used to support process execution.

No	Name	Description

Figure A.4 Process Attributes Description

			Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Metric Name:			F : Indicatiors of the atribute are fully satisfied (%86-100)		
Conceptual Definition:			L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed On:			P : Indicatiors of the atribute are largely satisfied (%16-50)		
Assessed By:			N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes			Answers	Rating	Expected Answers
	Indicators				
Metric Identity			MUF-1	F	
	Q1	Which entity does the metric measure?			
	Q2	Which attribute of the entity does the metric measure?			
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)			Ratio, Absolute
	Q4	What is the unit of the metric data?			
	Q5	What is the type of the metric data? (integer, real, etc.)			
	Q6	What is the range of the metric data?			
Data Existence			MUF-2	F	
	Q7	Is metric data existent?			Available > 20
	Q8	What is the amount of overall observations?			
	Q9	What is the amount of missing data points?			
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)			
	Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)			
Data Verifiability			MUF-3	F	
	Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)			
	Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)			Yes
	Q14	Who is responsible for recording metric data?			
	Q15	Is all metric data recorded by the responsible body?			Yes
	Q16	How is metric data recorded? (on a form, report, tool, etc.)			
	Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)			Yes
	Q18	Where is metric data stored? (in a file, database, etc.)			
	Q19	Is all metric data stored in the same place? (in a file, database, etc.)			Yes
Data Dependability			MUF-4	F	
	Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)			
	Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)			
	Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)			
	Q23	Are the frequencies for data generation, recording, and storing different?			No
	Q24	Is metric data recorded precisely?			Yes
	Q25	Is metric data collected for a specific purpose?			Yes
	Q26	Is the purpose of metric data collection known by process performers?			Yes
	Q27	Is metric data analyzed and reported?			Yes
	Q28	Is metric data analysis results communicated to process performers?			Yes
	Q29	Is metric data analysis results communicated to management?			Yes
	Q30	Is metric data analysis results used as a basis for decision making?			Yes
Data Normalizability					
	Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)			
Data Integrability					
	Q32	Is metric data integrable at project level?			
	Q33	Is metric data integrable at organization level?			

Metric Name:		
Conceptual Definition:		
Assessed On:		
Assessed By:		
Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	F	L or F
Metric Usability Result	**U**	L or F (Usable) -- Not Usable otherwise

Figure A.5 Metric Usability Questionnaire for Base Metrics

			Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Metric Name:			F : Indicatiors of the atribute are fully satisfied (%86-100)		
Conceptual Definition:			L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed On:			P : Indicatiors of the atribute are largely satisfied (%16-50)		
Assessed By:			N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes			**Answers**	**Rating**	Expected Answers
	Indicators				
Metric Identity			**MUF-1**	**F**	
	Q1	What is the the metric formula? (please refer to related base metrics)			
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)			Ratio, Absolute
	Q3	What is the unit of the metric data?			
	Q4	What is the type of the metric data? (integer, real, etc.)			
	Q5	What is the range of the metric data?			
Data Existence			**MUF-2**	**F**	
	Q6	Is metric data existent?			Available > 10
	Q7	What is the amount of overall observations?			
	Q8	What is the amount of missing data points?			
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)			
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)			
Data Verifiability			**MUF-3**	**F**	
	Q11	How is metric data calculated? (by a tool, manually, etc.)			
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)			Yes
	Q13	Is all metric data calculated according to metric formula?			Yes
	Q14	Where is metric data stored? (in a file, database, etc.)			
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)			Yes
Data Dependability			**MUF-4**	**F**	
	Q16	Is metric data stored precisely?			Yes
	Q17	Is metric data stored for a specific purpose?			Yes
	Q18	Is the purpose of metric data storage known by process performers?			Yes
	Q19	Is metric data analyzed and reported?			Yes
	Q20	Is metric data analysis results communicated to process performers?			Yes
	Q21	Is metric data analysis results communicated to management?			Yes
	Q22	Is metric data analysis results used as a basis for decision making?			Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)			
Data Integrability					
	Q24	Is metric data integrable at project level?			
	Q25	Is metric data integrable at organization level?			

Metric Name:		
Conceptual Definition:		
Assessed On:		
Assessed By:		
Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	F	L or F
MUF-3&4 for base metric-1		L or F
MUF-3&4 for base metric-2		L or F
MUF-3&4 for base metric-n		L or F
Metric Usability Result	**U**	L or F (Usable) -- Not Usable otherwise

Figure A.6 Metric Usability Questionnaire for Derived Metrics

APPENDIX B. CASE STUDY PLANS

Work Plan for Case Study-A

We had a meeting with the Team Leader of the project, and explained the purpose and context of the work. We then prepared a work plan as shown in Figure B.1. We estimated the effort required by us and by the company separately, and summed these values under the column of planned total effort. We also recorded actual effort spent by us and by the company, and summed them under the column of actual total effort. We spent 42 man-hours for exporting and reviewing task management data, applying the approach, performing the analyses, and interpreting the results. We used 4 man-hours of the Team Leader. Overall, we spent 46 man-hours for the case study implementation (the estimation was 33 man-hours).

Task No	Task Name	Duration (days)	Start	Finish	Planned Work (by us) (hours)	Planned Work (by company) (hours)	**Planned Total Work (hours)**	Actual Work (by us) (hours)	Actual Work (by company) (hours)	**Actual Total Work (hours)**	Pred. Task No
1	Task Management Process Case Study	13,0	03.06.2005	15.06.2005	30	3	**33**	42	4	**46**	
2	Meet with Team Leader to understand the process and data	1,0	03.06.2005	03.06.2005	1	1	**2**	1	1	**2**	
3	Export task management data into a file, and review data	1,0	03.06.2005	03.06.2005	2	0	**2**	4	1	**5**	
4	Perform rational sampling of process executions and data	3,0	07.06.2005	09.06.2005	8	0	**8**	11	0	**11**	3
5	Evaluate the utilization of metrics	1,0	07.06.2005	07.06.2005	2	0	**2**	3	0	**3**	3
6	Identify resulting process clusters and metrics for control charting	1,0	10.06.2005	10.06.2005	1	0	**1**	2	0	**2**	4;5
7	Perform control charting and interpret initial results	2,0	10.06.2005	11.06.2005	10	0	**10**	14	0	**14**	6
8	Interview with Team Leader to understand assignable causes	1,0	12.06.2006	12.06.2006	2	2	**4**	2	2	**4**	7
9	Remove assignable causes, re-chart data, and interpret the results	3,0	13.06.2005	15.06.2005	4	0	**4**	5	0	**5**	8

Figure B.1 Work Plan for Case Study-A

Work Plan for Case Study-B

We planned case study implementation after holding a meeting with the Quality Assurance Expert of the company. She had contributed to all reviews and she had been the owner of the review process since its establishment in 2004. In the meeting, we explained the aim and the context of the work that we intended to perform, and requested her contribution at certain points. After her commitment, we prepared a work plan as shown in Figure B.2. We estimated the effort required by us and by the company separately, and showed their sum in the column of planned total effort. Similarly, we recorded actual effort spent by us and by the company, and depicted their sum in the column of actual total effort. We spent 95 man-hours for gathering and translating review data, applying the approach, performing the analyses, and interpreting the results. We used 20 man-hours of the company staff. Overall, case study implementation took 115 man-hours, though our estimation was 100 man-hours.

Task No	Task Name	Duration (days)	Start	Finish	Planned Work (by us) (hours)	Planned Work (by company) (hours)	**Planned Total Work (hours)**	Actual Work (by us) (hours)	Actual Work (by company) (hours)	**Actual Total Work (hours)**	Pred. Task No
1	Review Process Case Study	91,0	26.08.2005	24.11.2005	82	18	**100**	95	20	**115**	
2	Meet with QA Expert for study plan	1,0	26.08.2005	26.08.2005	2	2	**4**	2	2	**4**	
3	Plan the case study	1,0	01.09.2005	01.09.2005	6	0	**6**	9	0	**9**	2
4	Review process records and gather process data in a file	2,0	21.09.2005	22.09.2005	12	6	**18**	11	4	**15**	3
5	Perform rational sampling of process executions and data	9,0	27.09.2005	05.10.2005	24	0	**24**	32	0	**32**	4
6	Evaluate the utilization of review metrics	1,0	11.10.2005	11.10.2005	8	0	**2**	6	0	**6**	5
7	Identify resulting process clusters and process metrics as basis for control charting	1,0	14.10.2005	14.10.2005	2	0	**2**	4	0	**4**	6;5
8	Perform control charting and interpret initial results	8,0	18.10.2005	25.10.2005	16	0	**16**	19	0	**19**	
9	for nonconformance detection efficiency	1,0	18.10.2005	18.10.2005	4	0	**4**	6	0	**6**	7
10	for nonconformance resolution efficiency	1,0	20.10.2005	20.10.2005	4	0	**4**	5	0	**5**	9
11	for open period	1,0	24.10.2005	24.10.2005	4	0	**4**	3	0	**3**	10
12	for open period with respect to nonconformances	1,0	25.10.2005	25.10.2005	4	0	**4**	5	0	**5**	11
13	Interview with QA Expert to understand any assignable causes	1,0	03.11.2005	03.11.2005	2	2	**4**	2	2	**4**	12
14	Remove assignable causes, re-chart data, and interpret the results	2,0	08.11.2005	09.11.2005	8	0	**8**	8	0	**8**	
15	for nonconformance detection efficiency	1,0	08.11.2005	08.11.2005	2	0	**2**	3	0	**3**	13
16	for nonconformance resolution efficiency	1,0	08.11.2005	08.11.2005	2	0	**2**	2	0	**2**	15
17	for open period	1,0	09.11.2005	09.11.2005	2	0	**2**	1	0	**1**	16
18	for open period with respect to nonconformances	1,0	09.11.2005	09.11.2005	2	0	**2**	2	0	**2**	17
19	Share analysis results with company staff in a meeting	1,0	24.11.2005	24.11.2005	2	8	**10**	2	12	**14**	18

Figure B.2 Work Plan for Case Study-B

Work Plan for Case Studies C, D, and E

We planned case study implementation after holding a meeting with a Team Leader of test development teams. He led two test development teams of the project since 2003. In the meeting, we explained the aim and the context of the work, and requested his contribution. We decided on the processes to work with, based on process relations and availability of process data. Accordingly we prepared a work plan as shown in Figure B.3. We estimated the effort required and recorded actual effort spent by us and by the company separately. We also summarized total planned and actual effort in separate columns. We spent 36 man-hours for gathering and translating review data, applying the approach, performing the analyses, and interpreting the results. We used 8 man-hours of the company staff. Overall, case study implementation took 44 man-hours, though our estimation was 54 man-hours.

Task No	Task Name	Duration (days)	Start	Finish	Planned Work (by us) (hours)	Planned Work (by company) (hours)	**Planned Total Work (hours)**	Actual Work (by us) (hours)	Actual Work (by company) (hours)	**Actual Total Work (hours)**	Pred. Task No
1	Test Development Case Study	86,0	10.03.2006	03.06.2006	42	12	54	36	8	44	
2	Meet with TL for study plan	1,0	10.03.2006	10.03.2006	1	1	**2**	1	1	**2**	
3	Plan the case study	1,0	11.03.2006	11.03.2006	2	0	**2**	2	0	**2**	2
4	Depict general process flows	2,0	20.03.2006	21.03.2006	4	2	**6**	3	1	**4**	3
5	Review process data, and generate data sets to work on	37,0	23.03.2006	28.04.2006	6	2	**8**	6	2	**8**	3
6	Perform rational sampling of process executions and data	2,0	29.04.2006	30.04.2006	3	0	**3**	3	0	**3**	5
7	Generate derived data sets	1,0	01.05.2006	01.05.2006	4	2	**6**	2	1	**3**	5
8	Evaluate the utilization of process metrics	3,0	02.05.2006	04.05.2006	6	0	**6**	3	0	**3**	5
9	Identify resulting process clusters and process metrics as basis for control charting	1,0	05.05.2006	05.05.2006	2	0	**2**	1	0	**1**	6;7;8
10	Perform control charting and interpret initial results	5,0	05.05.2006	09.05.2006	8	2	**10**	7	0	**7**	9
11	Interview with TL to understand any assignable causes	1,0	01.06.2006	01.06.2006	2	2	**4**	2	2	**4**	10
12	Remove assignable causes, re-chart data, and interpret the results	2,0	02.06.2006	03.06.2006	4	1	**5**	6	1	**7**	11

Figure B.3 Work Plan for Case Studies C, D, and E

Work Breakdown Structure

Detailed description of the work as a result of case study implementations is provided in Table B.1 as a work breakdown structure.

Table B.1 Work Breakdown Structure for Case Studies

1. Review process records and gather process data in a file
   - Consolidate review data in time sequence and in a form that is appropriate for comparison among different projects and product types
   - During consolidation, establish traceability between process executions and review data
   - Eliminate any missing, incomplete, or invalid data points
2. Perform rational sampling of process executions
   - Sample 3-5 review records and fill Process Execution Record for each
   - Obtain a merged list of process attribute values from sample Process Execution Records
   - Record the merged list of process attribute values into the rows of Process Similarity Matrix
   - Record the numbers of entire process executions into the columns of Process Similarity Matrix
   - Verify each column of Process Similarity Matrix against process attribute values recorded by rows
   - During verification, if a process execution has a process attribute value out of recorded ones, add that process attribute value as a row under its process attribute category within the matrix
   - When verification is completed, review the Process Similarity Matrix by columns and take a copy of each column if it is different from previous columns in terms of process attribute values
   - Label each copied column as a process cluster
   - Calculate distances (number of differing attribute values) between process clusters
   - Identify possible merges among the clusters that have a maximum distance of 1 between them
   - Identify initial process clusters
3. Evaluate the utilization of review metrics
   - Answer Metric Usability Questionnaire for each base metric
   - Answer Metric Usability Questionnaire for each derived metric
   - Identify the usability of base and derived metrics according to related metric usability indices
4. Identify resulting process clusters and process metrics as basis for control charting
   - Review initial process clusters
   - Review usability of process metrics
   - Review process data for each process cluster-process metric pair
   - Eliminate process cluster-process metric pairs for which the data is not available
   - Identify resulting process cluster-process metric pairs as basis for control charting
5. For each process cluster-process metric pair, perform the following:
   - Transfer the data into statistical analysis tool, and remove any missing data points
   - Chart the data and interpret initial results
6. Understand the assignable causes, if any, regarding control charts
   - Interview with process performers for any potential assignable causes
   - Answer Process Execution Questionnaire for each process execution regarding an out-of-control point to understand the assignable causes
7. For each control chart, perform the following:
   - Remove the assignable causes, and re-chart the data
   - Interpret the results from re-charted data

APPENDIX C. DETAILS OF CASE STUDY-A

SPC-AM Assets

Process Execution Records

SPC[AM]
Process Execution Record
(Internal Attributes)

Process Name:	Task Management	Recorded On:	7.June.2005
Process Execution No:	5	Recorded By:	AT

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	Task request	

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Analysis knowledge	"Intersept'lerin Analizi (ELDES Interseptleri)"

3. Activities (in sequence): Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Enter task request	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Task assigner	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
1	Starteam	

Figure C.1 Process Execution Record for Task Management PE # 5

SPC[TM]
Process Execution Record
(Internal Attributes)

Process Name:	Task Management	Recorded On:	7.June.2005
Process Execution No:	10	Recorded By:	AT

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	Task request	

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Document	"Detayli Tasarim Dokumaninin (SDD) Tamamlanmasi"

3. Activities (in sequence): Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Enter task request	
2	Implement task request	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Task assigner	
2	Task implementer	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
1	Starteam	

Figure C.2 Process Execution Record for Task Management PE # 10

SPC[TM]
Process Execution Record
(Internal Attributes)

Process Name:	Task Management	Recorded On:	7.June.2005
Process Execution No:	25	Recorded By:	AT

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	Task request	

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Software code	"Data Management paketine Transaction Manager class'inin tanimlanmasi"

3. Activities (in sequence): Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Enter task request	
2	Implement task request	
3	Verify implementation	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Task assigner	
2	Task implementer	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
1	Starteam	

Figure C.3 Process Execution Record for Task Management PE # 25

SPC[TM]
Process Execution Record
(Internal Attributes)

Process Name:	Task Management	Recorded On:	7.June.2005
Process Execution No:	57	Recorded By:	AT

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	Task request	

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Design	"Kutuphane Olusturma Modulu Tasarım Çalışması"

3. Activities (in sequence): Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Enter task request	
2	Implement task request	
3	Verify task request	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Task assigner	
2	Task implementer	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
1	Starteam	

Figure C.4 Process Execution Record for Task Management PE # 57

		Process Executions																						
Process Attributes		PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	PE21	PE22	PE23
1	**Inputs**																							
1.1	Task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2	**Outputs**																							
2.1	Document	•								•	•	•										•		
2.2	Software code		•	•										•	•	•								•
2.3	Analysis knowledge					•	•										•		•	•	•			
2.4	Design							•	•									•						
2.5	Research knowledge				•																		•	
2.6	Unclassified output												•											
3	**Activities**																							
3.1	Enter task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2	Implement task request	•	•	•	•		•		•	•	•	•	•			•	•	•	•	•	•	•	•	•
3.3	Verify task request				•				•							•			•		•			•
4	**Roles**																							
4.1	Task assigner	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2	Task implementer	•	•	•	•		•	•	•	•	•	•	•			•	•	•	•	•	•	•	•	•
5	**Tools and Techniques**																							
5.1	Starteam	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Process Cluster		C	C	C	D	A	C	B	D	C	C	C	C	A	A	D	C	C	D	C	D	C	C	D
Process Sub-cluster		1	2	2	5	3	3	4	4	1	1	1	6	2	2	2	3	4	3	3	3	1	5	2

		Process Executions																						
Process Attributes		PE24	PE25	PE26	PE27	PE28	PE29	PE30	PE31	PE32	PE33	PE34	PE35	PE36	PE37	PE38	PE39	PE40	PE41	PE42	PE43	PE44	PE45	PE46
1	**Inputs**																							
1.1	Task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2	**Outputs**																							
2.1	Document																							•
2.2	Software code		•	•	•				•	•				•	•	•	•	•		•	•	•	•	
2.3	Analysis knowledge							•											•					
2.4	Design						•																	
2.5	Research knowledge	•									•	•	•											
2.6	Unclassified output					•																		
3	**Activities**																							
3.1	Enter task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2	Implement task request	•	•	•	•		•			•	•	•	•			•		•	•					•
3.3	Verify task request						•			•	•	•						•	•					
4	**Roles**																							
4.1	Task assigner	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2	Task implementer	•	•	•	•	•	•	•	•	•	•	•	•			•		•	•					•
5	**Tools and Techniques**																							
5.1	Starteam	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Process Cluster		C	C	C	C	B	D	B	B	D	D	D	C	A	A	C	A	D	D	A	A	A	A	C
Process Sub-cluster		5	2	2	2	6	4	3	2	2	5	5	5	2	2	2	2	2	3	2	2	2	2	1

		Process Executions																						
Process Attributes		PE47	PE48	PE49	PE50	PE51	PE52	PE53	PE54	PE55	PE56	PE57	PE58	PE59	PE60	PE61	PE62	PE63	PE64	PE65	PE66	PE67	PE68	PE69
1	**Inputs**																							
1.1	Task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2	**Outputs**																							
2.1	Document			•																				
2.2	Software code	•	•		•	•	•			•	•			•		•	•	•	•	•	•	•	•	•
2.3	Analysis knowledge							•																
2.4	Design								•			•												
2.5	Research knowledge												•											
2.6	Unclassified output														•									
3	**Activities**																							
3.1	Enter task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2	Implement task request	•		•	•	•			•	•		•	•	•	•			•	•		•	•	•	•
3.3	Verify task request								•	•		•	•	•	•			•	•		•	•		
4	**Roles**																							
4.1	Task assigner	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2	Task implementer	•		•	•	•		•	•	•		•	•	•	•	•	•	•	•		•	•	•	•
5	**Tools and Techniques**																							
5.1	Starteam	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Process Cluster		C	A	C	C	C	A	B	D	D	A	D	D	D	D	B	B	D	D	A	D	D	C	C
Process Sub-cluster		2	2	1	2	2	2	3	4	2	2	4	5	2	6	2	2	2	2	2	2	2	2	2

		Process Executions																					
Process Attributes		PE70	PE71	PE72	PE73	PE74	PE75	PE76	PE77	PE78	PE79	PE80	PE81	PE82	PE83	PE84	PE85	PE86	PE87	PE88	PE89	PE90	PE91
1	**Inputs**																						
1.1	Task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2	**Outputs**																						
2.1	Document										•												
2.2	Software code	•	•		•		•	•	•	•		•	•	•	•	•	•	•	•			•	•
2.3	Analysis knowledge																				•		
2.4	Design																						
2.5	Research knowledge			•		•																	
2.6	Unclassified output																			•			
3	**Activities**																						
3.1	Enter task request	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2	Implement task request	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•		
3.3	Verify task request				•	•			•		•	•											
4	**Roles**																						
4.1	Task assigner	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2	Task implementer	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		
5	**Tools and Techniques**																						
5.1	Starteam	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Process Cluster		C	C	C	D	D	C	C	D	C	D	D	B	C	C	C	C	C	C	C	C	A	A
Process Sub-cluster		2	2	5	2	5	2	2	2	2	1	2	2	2	2	2	2	2	2	6	3	2	2

Figure C.5 Process Similarity Matrix for Task Management Process

Metric Usability Questionnaires

Metric Name: Estimated Task Start Date			**Assessed On:** 7.June.2005
Conceptual Definition:			**Assessed By:** AT
Attributes			**Answers**
Metric Definition			
	Q1	What is the name of the metric?	Estimated Task Start Date
	Q2	Which entity does the metric measure?	Task
	Q3	Which attribute of the entity does the metric measure?	Estimated start date
	Q4	What is the type of the metric? (direct, indirect)	Direct
	Q5	How is metric data calculated? (specify metric formula if the type is indirect)	N/A
	Q6	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Interval
	Q7	What is the unit of the metric data?	Date
	Q8	What is the type of the metric data? (integer, real, etc.)	Date
	Q9	What is the range of the metric data?	[25.Nov.2002, 22.Apr.2004]
Data Existence			
	Q10	Is metric data existent?	Yes
	Q11	What is the amount of metric data points?	67
	Q12	Are there any missing data points?	No
	Q13	What is the amount of missing data points?	0
Data Verifiability			
	Q14	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At start (while entering task request)
	Q15	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes
	Q16	Who is responsible for recording metric data?	Task assigner
	Q17	Is all metric data recorded by the responsible body?	Yes
	Q18	How is metric data recorded? (on a form, report, tool, etc.)	On a tool (Starteam)
	Q19	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes
	Q20	Where is metric data stored? (in a file, database, etc.)	Starteam database
	Q21	Is all metric data stored in the same place? (in a file, database, etc.)	Yes
Data Dependability			
	Q22	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is entered)
	Q23	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is entered)
	Q24	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is entered)
	Q25	Is metric data recorded precisely?	Yes
	Q26	Is metric data collected for a specific purpose?	Yes (for the purpose of task monitoring)
	Q27	Is the purpose of metric data collection known by process performers?	Yes
	Q28	Is metric data analyzed and reported?	No
	Q29	Is metric data analysis results communicated to process performers?	No
	Q30	Is metric data analysis results communicated to management?	No
	Q31	Is metric data analysis results used as a basis for decision making?	No
Data Normalizability			
	Q32	Is metric data normalizable per product type basis?	No
	Q33	Is metric data normalizable per process phase basis?	N/A
	Q34	Is metric data normalizable per resource type basis?	Yes
Data Integrability			
	Q35	Is metric data integrable at project level?	Yes
	Q36	Is metric data integrable at process level?	N/A
	Q37	Is metric data integrable at organization level?	N/A

Figure C.6 Metric Usability Questionnaire for "Estimated Start Date" Base Metric of Task Management Process

Metric Name: Estimated Task Finish Date			**Assessed On:** 7.June.2005
Conceptual Definition:			**Assessed By:** AT
Attributes			**Answers**
Metric Definition			
	Q1	What is the name of the metric?	Estimated Task Finish Date
	Q2	Which entity does the metric measure?	Task
	Q3	Which attribute of the entity does the metric measure?	Estimated finish date
	Q4	What is the type of the metric? (direct, indirect)	Direct
	Q5	How is metric data calculated? (specify metric formula if the type is indirect)	N/A
	Q6	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Interval
	Q7	What is the unit of the metric data?	Date
	Q8	What is the type of the metric data? (integer, real, etc.)	Date
	Q9	What is the range of the metric data?	[27.Dec.2002, 26.Apr.2004]
Data Existence			
	Q10	Is metric data existent?	Yes
	Q11	What is the amount of metric data points?	67
	Q12	Are there any missing data points?	No
	Q13	What is the amount of missing data points?	0
Data Verifiability			
	Q14	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At start (while entering task request)
	Q15	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes
	Q16	Who is responsible for recording metric data?	Task assigner
	Q17	Is all metric data recorded by the responsible body?	Yes
	Q18	How is metric data recorded? (on a form, report, tool, etc.)	On a tool (Starteam)
	Q19	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes
	Q20	Where is metric data stored? (in a file, database, etc.)	Starteam database
	Q21	Is all metric data stored in the same place? (in a file, database, etc.)	Yes
Data Dependability			
	Q22	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is entered)
	Q23	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is entered)
	Q24	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is entered)
	Q25	Is metric data recorded precisely?	Yes
	Q26	Is metric data collected for a specific purpose?	Yes (for the purpose of task monitoring)
	Q27	Is the purpose of metric data collection known by process performers?	Yes
	Q28	Is metric data analyzed and reported?	No
	Q29	Is metric data analysis results communicated to process performers?	No
	Q30	Is metric data analysis results communicated to management?	No
	Q31	Is metric data analysis results used as a basis for decision making?	No
Data Normalizability			
	Q32	Is metric data normalizable per product type basis?	No
	Q33	Is metric data normalizable per process phase basis?	N/A
	Q34	Is metric data normalizable per resource type basis?	Yes
Data Integrability			
	Q35	Is metric data integrable at project level?	Yes
	Q36	Is metric data integrable at process level?	N/A
	Q37	Is metric data integrable at organization level?	N/A

Figure C.7 Metric Usability Questionnaire for “Estimated Finish Date” Base Metric of Task Management Process

Metric Name: Actual Task Start Date			**Assessed On:** 7.June.2005
Conceptual Definition:			**Assessed By:** AT
Attributes			**Answers**
Metric Definition			
	Q1	What is the name of the metric?	Actual Task Start Date
	Q2	Which entity does the metric measure?	Task
	Q3	Which attribute of the entity does the metric measure?	Actual start date
	Q4	What is the type of the metric? (direct, indirect)	Direct
	Q5	How is metric data calculated? (specify metric formula if the type is indirect)	N/A
	Q6	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Interval
	Q7	What is the unit of the metric data?	Date
	Q8	What is the type of the metric data? (integer, real, etc.)	Date
	Q9	What is the range of the metric data?	[3.Jan.2003, 26.Apr.2004]
Data Existence			
	Q10	Is metric data existent?	Yes
	Q11	What is the amount of metric data points?	67
	Q12	Are there any missing data points?	Yes
	Q13	What is the amount of missing data points?	27
Data Verifiability			
	Q14	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At middle (while accepting task request)
	Q15	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes
	Q16	Who is responsible for recording metric data?	Task implementer
	Q17	Is all metric data recorded by the responsible body?	Yes
	Q18	How is metric data recorded? (on a form, report, tool, etc.)	On a tool (Starteam)
	Q19	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes
	Q20	Where is metric data stored? (in a file, database, etc.)	Starteam database
	Q21	Is all metric data stored in the same place? (in a file, database, etc.)	Yes
Data Dependability			
	Q22	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is accepted)
	Q23	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is accepted)
	Q24	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when a new task request is accepted)
	Q25	Is metric data recorded precisely?	Yes
	Q26	Is metric data collected for a specific purpose?	Yes (for the purpose of task monitoring)
	Q27	Is the purpose of metric data collection known by process performers?	Yes
	Q28	Is metric data analyzed and reported?	No
	Q29	Is metric data analysis results communicated to process performers?	No
	Q30	Is metric data analysis results communicated to management?	No
	Q31	Is metric data analysis results used as a basis for decision making?	No
Data Normalizability			
	Q32	Is metric data normalizable per product type basis?	No
	Q33	Is metric data normalizable per process phase basis?	N/A
	Q34	Is metric data normalizable per resource type basis?	Yes
Data Integrability			
	Q35	Is metric data integrable at project level?	Yes
	Q36	Is metric data integrable at process level?	N/A
	Q37	Is metric data integrable at organization level?	N/A

Figure C.8 Metric Usability Questionnaire for “Actual Start Date” Base Metric of Task Management Process

Metric Name: Actual Task Finish Date			**Assessed On:** 7.June.2005
Conceptual Definition:			**Assessed By:** AT
Attributes			**Answers**
Metric Definition			
	Q1	What is the name of the metric?	Actual Task Finish Date
	Q2	Which entity does the metric measure?	Task
	Q3	Which attribute of the entity does the metric measure?	Actual finish date
	Q4	What is the type of the metric? (direct, indirect)	Direct
	Q5	How is metric data calculated? (specify metric formula if the type is indirect)	N/A
	Q6	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Interval
	Q7	What is the unit of the metric data?	Date
	Q8	What is the type of the metric data? (integer, real, etc.)	Date
	Q9	What is the range of the metric data?	[13.Jan.2003, 18.Oct.2004]
Data Existence			
	Q10	Is metric data existent?	Yes
	Q11	What is the amount of metric data points?	67
	Q12	Are there any missing data points?	No
	Q13	What is the amount of missing data points?	0
Data Verifiability			
	Q14	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end (when task request is finished)
	Q15	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes
	Q16	Who is responsible for recording metric data?	Task implementer
	Q17	Is all metric data recorded by the responsible body?	Yes
	Q18	How is metric data recorded? (on a form, report, tool, etc.)	On a tool (Starteam)
	Q19	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes
	Q20	Where is metric data stored? (in a file, database, etc.)	Starteam database
	Q21	Is all metric data stored in the same place? (in a file, database, etc.)	Yes
Data Dependability			
	Q22	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when task request is finished)
	Q23	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when task request is finished)
	Q24	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronously (when task request is finished)
	Q25	Is metric data recorded precisely?	Yes
	Q26	Is metric data collected for a specific purpose?	Yes (for the purpose of task monitoring)
	Q27	Is the purpose of metric data collection known by process performers?	Yes
	Q28	Is metric data analyzed and reported?	No
	Q29	Is metric data analysis results communicated to process performers?	No
	Q30	Is metric data analysis results communicated to management?	No
	Q31	Is metric data analysis results used as a basis for decision making?	No
Data Normalizability			
	Q32	Is metric data normalizable per product type basis?	No
	Q33	Is metric data normalizable per process phase basis?	N/A
	Q34	Is metric data normalizable per resource type basis?	Yes
Data Integrability			
	Q35	Is metric data integrable at project level?	Yes
	Q36	Is metric data integrable at process level?	N/A
	Q37	Is metric data integrable at organization level?	N/A

Figure C.9 Metric Usability Questionnaire for “Actual Finish Date” Base Metric of Task Management Process

Process Execution Questionnaires

Process Name: Task Management		Recorded On: 12.June.2005	
Process Execution No: 1		Recorded By: AT	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The work plan changed, and task assignment was not updated.		

Figure C.10 Process Execution Questionnaire for Task Management PE # 1

Process Name: Task Management		Recorded On: 12.June.2005	
Process Execution No: 26		Recorded By: AT	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	Development of support tool required a much deeper study and greater effort		

Figure C.11 Process Execution Questionnaire for Task Management PE # 26

Process Name: Task Management		Recorded On: 12.June.2005	
Process Execution No: 34		Recorded By: AT	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The task was forgotten to close.		

Figure C.12 Process Execution Questionnaire for Task Management PE # 34

Process Name: Task Management		Recorded On: 12.June.2005	
Process Execution No: 49		Recorded By: AT	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The task was forgotten to close.		

Figure C.13 Process Execution Questionnaire for Task Management PE # 49

Process Name: Task Management		Recorded On: 12.June.2005	
Process Execution No: 79		Recorded By: AT	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The task was forgotten to close.		

Figure C.14 Process Execution Questionnaire for Task Management PE # 79

Process Attributes Descriptions

SPC[tm]

Process Attributes Description

Process Name:	Task Management	Described On:	15 July 2005
Process Cluster:	C2	Described By:	AT

1. **Inputs:** Please list the inputs to the process.

No	Name	Description
1	Task request	

2. **Outputs:** Please list the outputs from the process.

No	Name	Description
1	Software code	

3. **Activities (in sequence):** Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Enter task request	
2	Implement task request	

4. **Roles:** Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Task assigner	
2	Task implementer	

5. **Tools and Techniques:** Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	Starteam	

Figure C.15 Process Attributes Description for Process Cluster C2

SPC

Process Attributes Description

Process Name:	Task Management	Described On:	15 July 2005
Process Cluster:	C-except C2	Described By:	AT

1. **Inputs:** Please list the inputs to the process.

No	Name	Description
1	Task request	

2. **Outputs:** Please list the outputs from the process.

No	Name	Description
1	Output type	Can have values: Document, Analysis Knowledge, Design, Research Knowledge, Unclassified Output

3. **Activities (in sequence):** Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Enter task request	
2	Implement task request	

4. **Roles:** Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Task assigner	
2	Task implementer	

5. **Tools and Techniques:** Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	Starteam	

Figure C.16 Process Attributes Description for Process Cluster C-except C2

SPC

Process Attributes Description

Process Name:	Task Management	Described On:	15 July 2005
Process Cluster:	D	Described By:	AT

1. **Inputs:** Please list the inputs to the process.

No	Name	Description
1	Task request	

2. **Outputs:** Please list the outputs from the process.

No	Name	Description
1	Output type	Can have values: Document, Software Code, Analysis Knowledge, Design, Research Knowledge, Unclassified Output

3. **Activities (in sequence):** Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Enter task request	
2	Implement task request	
3	Verify task request	

4. **Roles:** Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Task assigner	
2	Task implementer	

5. **Tools and Techniques:** Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	Starteam	

Figure C.17 Process Attributes Description for Process Cluster D

Control Charts

Estimation Capability

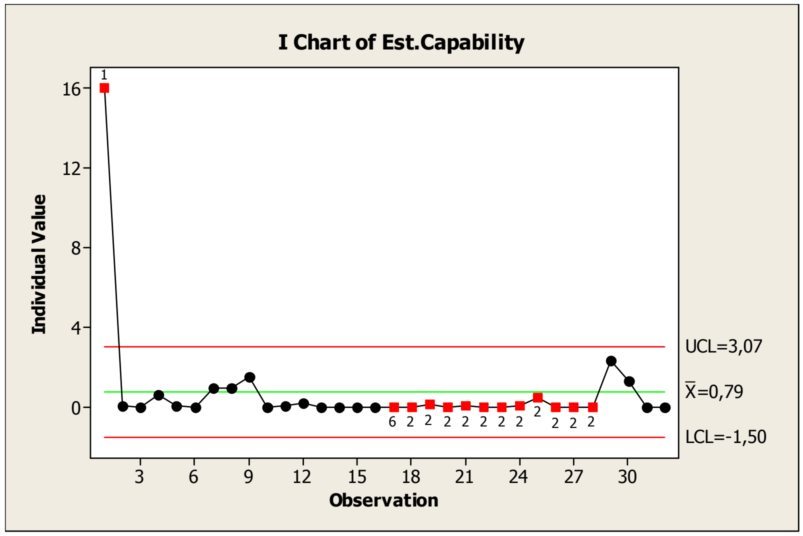


Figure C.18 Control Chart for Estimation Capability of Overall Task Management Process

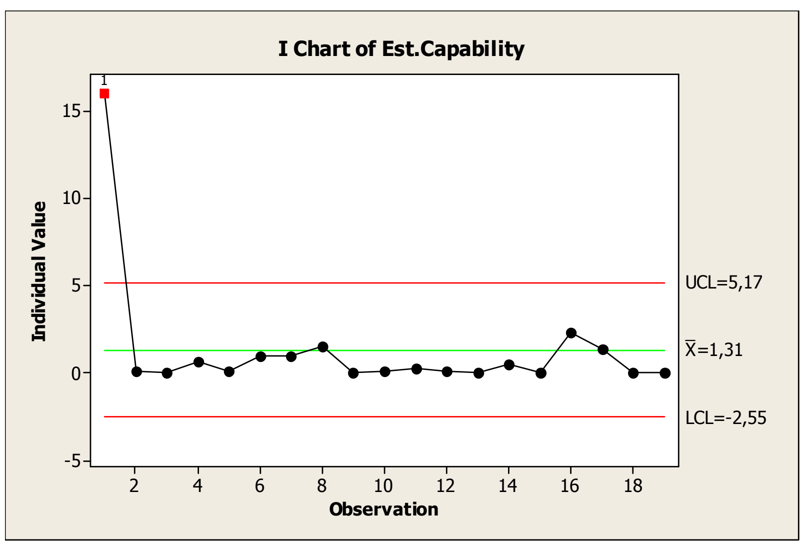


Figure C.19 Control Chart for Estimation Capability of Task Management Process Cluster C

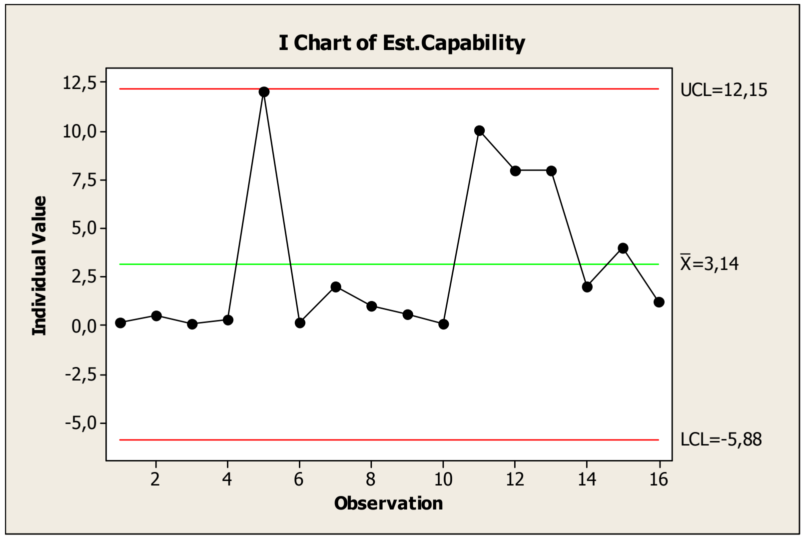


Figure C.20 Control Chart for Estimation Capability of Task Management Process Cluster C2

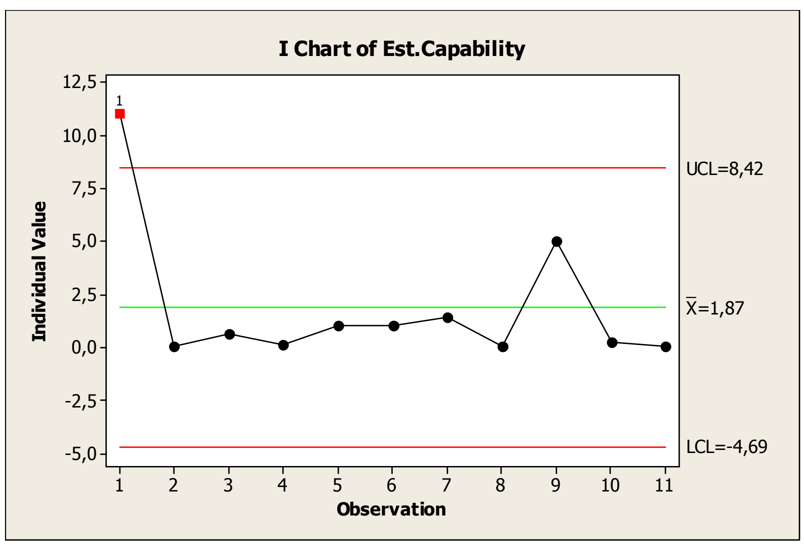


Figure C.21 Control Chart for Estimation Capability of Task Management Process Cluster C-except C2

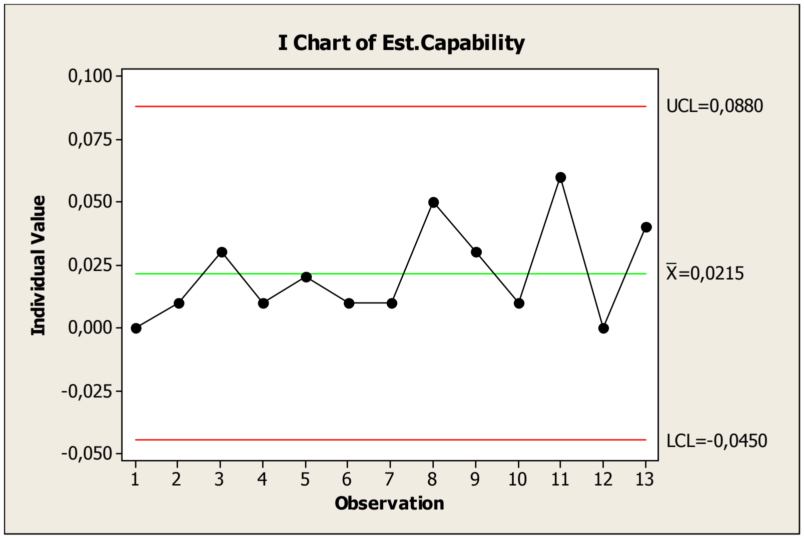


Figure C.22 Control Chart for Estimation Capability of Task Management Process Cluster D

Effort Variance

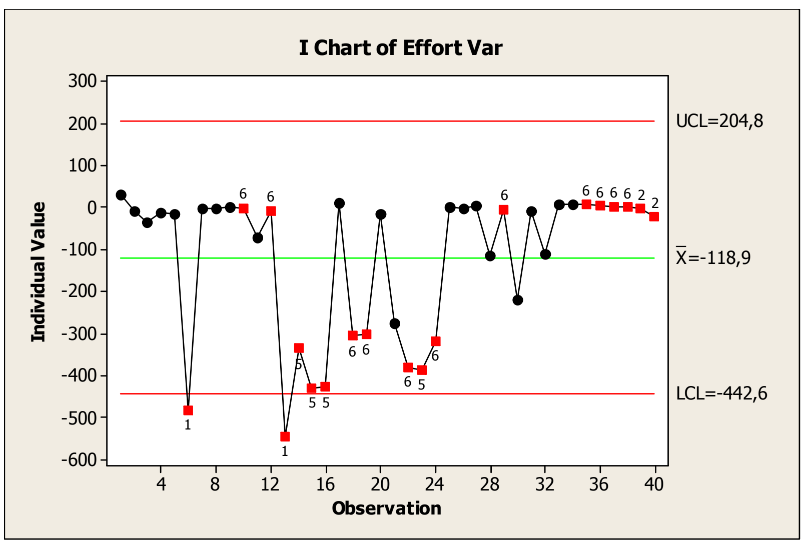


Figure C.23 Control Chart for Effort Variance of Overall Task Management Process

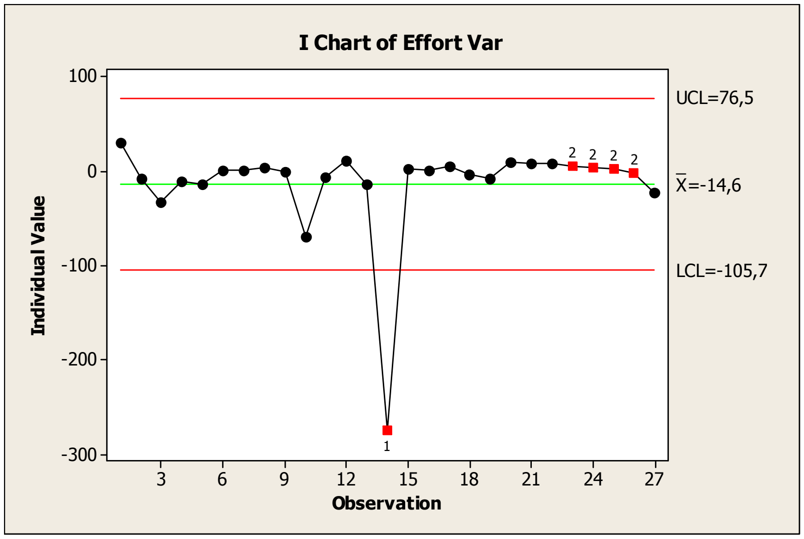


Figure C.24 Control Chart for Effort Variance of Task Management Process Cluster C

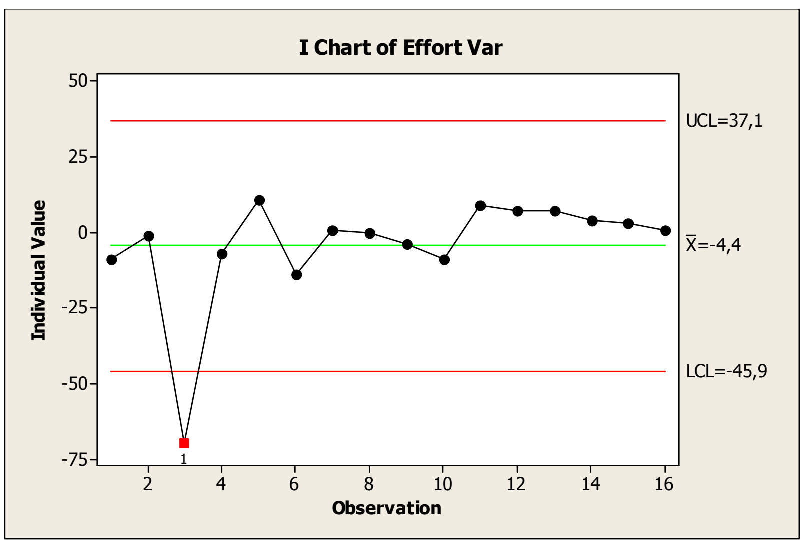


Figure C.25 Control Chart for Effort Variance of Task Management Process Cluster C2

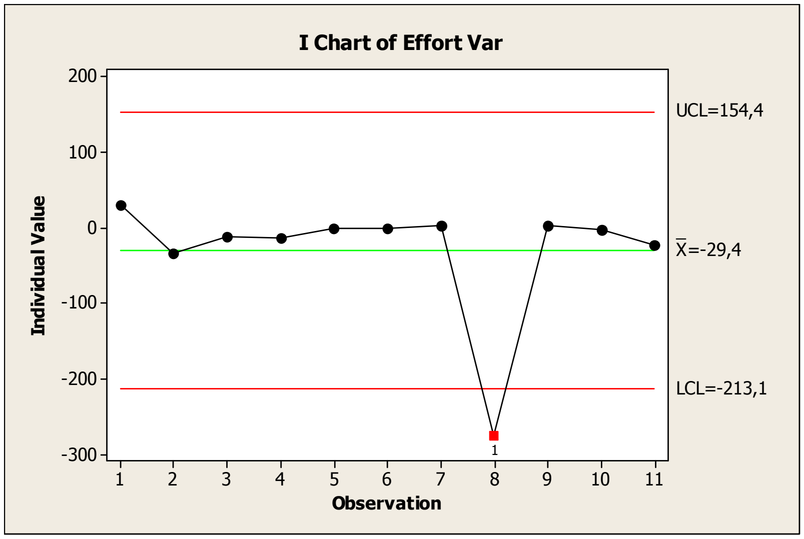


Figure C.26 Control Chart for Effort Variance of Task Management Process Cluster C-except C2

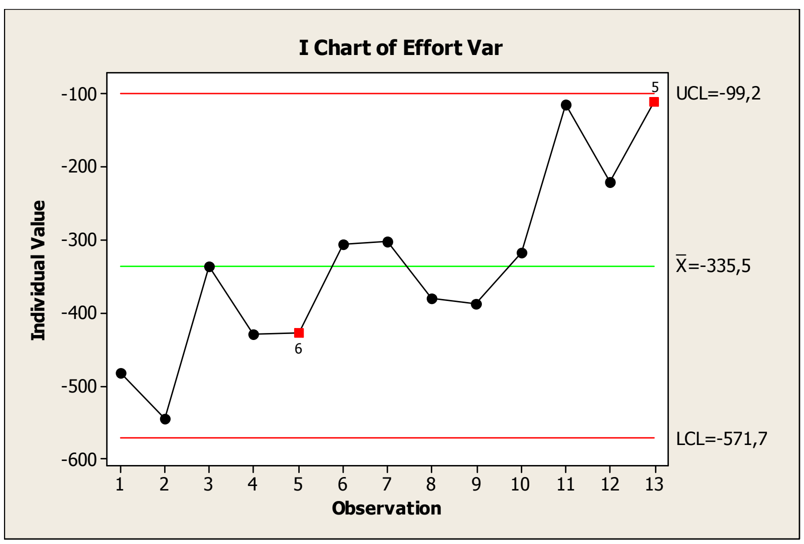
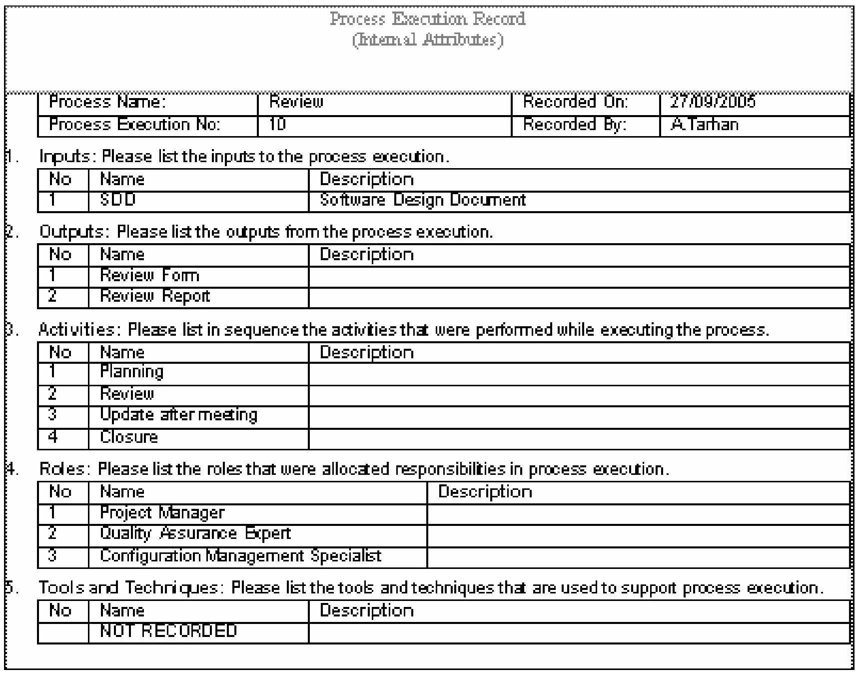


Figure C.27 Control Chart for Effort Variance of Task Management Process Cluster D

APPENDIX D. DETAILS OF CASE STUDY-B

SPC-AM Assets

Process Execution Records



Process Execution Record
(Internal Attributes)

Process Name:	Review	Recorded On:	27/09/2005
Process Execution No:	10	Recorded By:	A.Tarhan

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	SDD	Software Design Document

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Review Form	
2	Review Report	

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Planning	
2	Review	
3	Update after meeting	
4	Closure	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
	NOT RECORDED	

Figure D.1 Process Execution Record for Review PE # 10

Process Execution Record
(Internal Attributes)

Process Name:	Review	Recorded On:	27/09/2005
Process Execution No:	30	Recorded By:	A.Tarhan

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	YGD	Yazılım Gereksinimleri Dokümanı

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Review Report	

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Planning	
2	Review	
3	Closure	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
	NOT RECORDED	

Figure D.2 Process Execution Record for Review PE # 30

Process Execution Record
(Internal Attributes)

Process Name:	Review	Recorded On:	27/09/2005
Process Execution No:	53	Recorded By:	A.Tarhan

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	YTTD	Yazılım Test Tanımları Dokümanı

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Review Form	
2	Review Report	

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Planning	
2	Review	
3	Closure	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	
4	Customer	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
	NOT RECORDED	

Figure D.3 Process Execution Record for Review PE # 53

Process Execution Record
(Internal Attributes)

Process Name:	Review	Recorded On:	27/09/2005
Process Execution No:	177	Recorded By:	A.Tarhan

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	SSS	System Subsystem Specification

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Review Form	
2	Review Report	

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Planning	
2	Review	
3	Update after meeting	
4	Closure	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	
4	Customer	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
	NOT RECORDED	

Figure D.4 Process Execution Record for Review PE # 177

Process Execution Record
(Internal Attributes)

Process Name:	Review	Recorded On:	27/09/2005
Process Execution No:	189	Recorded By:	A.Tarhan

1. Inputs: Please list the inputs to the process execution.

No	Name	Description
1	YTR	Yazılım Test Raporu

2. Outputs: Please list the outputs from the process execution.

No	Name	Description
1	Review Form	
2	Review Report	

3. Activities: Please list in sequence the activities that were performed while executing the process.

No	Name	Description
1	Planning	
2	Review	
3	Update during meeting	
4	Closure	

4. Roles: Please list the roles that were allocated responsibilities in process execution.

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	

5. Tools and Techniques: Please list the tools and techniques that are used to support process execution.

No	Name	Description
	NOT RECORDED	

Figure D.5 Process Execution Record for Review PE # 189

Process Attributes	PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	PE21	PE22	PE23
1 Inputs																							
Product to review	SPS	SRS	STP	STD	SDD	DBDD	SRS	STP	STD	SDD	STP	SDP	STD	SRS	YGD	YGD	AKD1	AKD1	VYGD	AKD2	AKD3	AKD4	AKD2
1.1 Product type	4	4	1	2	2	2	4	1	2	2	1	1	2	4	4	4	2	2	8	2	2	2	2
2 Outputs																							
2.1 Review form	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•
2.2 Review report	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2.3 Code review report																							
3 Activities																							
3.1 Planning	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2 Review	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.3 Update during meeting																							
3.4 Update after meeting	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•
3.5 Closure	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4 Roles																							
4.1 Project Manager	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2 QA Expert	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.3 CM Specialist	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.4 Customer																							
5 Tools and Techniques																							
NOT IDENTIFIABLE																							
Process Cluster	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	A	A	A	A
Process Sub-cluster	4	4	1	2	2	2	4	1	2	2	1	1	2	4	4	4	2	2	8	2	2	2	2

Process Attributes	PE24	PE25	PE26	PE27	PE28	PE29	PE30	PE31	PE32	PE33	PE34	PE35	PE36	PE37	PE38	PE39	PE40	PE41	PE42	PE43	PE44	PE45	PE46
1 Inputs																							
Product to review	AKD3	AKD4	YGP	PYP	KYP	YKP	YGD	YGD	AKD1	AKD1	AKD2	AKD3	AKD4	AKD2	AKD3	AKD4	YGP	PYP	KYP	YKP	SDD-KMM	OEG	OEG
1.1 Product type	2	2	1	1	1	1	4	4	2	2	2	2	2	2	2	2	1	1	1	1	2	4	4
2 Outputs																							
2.1 Review form	•	•	•	•	•	•																•	•
2.2 Review report	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
2.3 Code review report																							
3 Activities																							
3.1 Planning	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2 Review	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.3 Update during meeting																							
3.4 Update after meeting	•	•	•	•	•	•																•	•
3.5 Closure	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4 Roles																							
4.1 Project Manager	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2 QA Expert	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.3 CM Specialist	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.4 Customer																							
5 Tools and Techniques																							
NOT IDENTIFIABLE																							
Process Cluster	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B	B	B	B	B	A	A
Process Sub-cluster	2	2	1	1	1	1	4	4	2	2	2	2	2	2	2	2	1	1	1	1	2	4	4

Process Attributes	PE47	PE48	PE49	PE50	PE51	PE52	PE53	PE54	PE55	PE56	PE57	PE58	PE59	PE60	PE61	PE62	PE63	PE64	PE65	PE66	PE67	PE68	PE69	PE70	PE71	PE72	PE73
1 Inputs																											
Product to review	OEG	OEG	SDD	SDD	SDD-KMM	SDD-KMM	YTTD	YTTD	YTTD	YTTD	TP	YTTD	KYP	AKD4	YTTD	YKP	YGP	PYP	SDD	AKD4	TP	YTTD	QTP	QTPR	YTP	KOD	YTP
1.1 Product type	4	4	2	2	2	2	5	5	5	5	3	5	1	2	5	1	1	1	2	2	3	5	1	5	1	3	1
2 Outputs																											
2.1 Review form					•		•	•			•		•	•		•	•	•	•	•			•	•	•		
2.2 Review report	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		•
2.3 Code review report																										•	
3 Activities																											
3.1 Planning	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2 Review	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.3 Update during meeting																											
3.4 Update after meeting					•						•		•	•		•	•	•	•	•			•	•	•		
3.5 Closure	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4 Roles																											
4.1 Project Manager	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2 QA Expert	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.3 CM Specialist	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.4 Customer							•	•																			
5 Tools and Techniques																											
NOT IDENTIFIABLE																											
Process Cluster	B	B	B	B	A	B	H	H	B	B	A	B	A	A	B	A	A	A	A	A	B	B	A	A	A	D	B
Process Sub-cluster	4	4	2	2	2	2	5	5	5	5	3	5	1	2	5	1	1	1	2	2	3	5	1	5	1	3	1

Process Attributes	PE74	PE75	PE76	PE77	PE78	PE79	PE80	PE81	PE82	PE83	PE84	PE85	PE86	PE87	PE88	PE89	PE90	PE91	PE92	PE93	PE94	PE95	PE96	PE97	PE98	PE99
1 Inputs																										
Product to review	QTPR	QTP	QTP	YTP	SPS	SVD	STD	SRS	STP	SDD	DBDD	KOD	YTP	SPS	STP	STD	KYP	SPS	STD	STR	STR	KOD	STD	SPS	SVD	STD
1.1 Product type	5	1	1	1	4	7	2	4	1	2	2	3	1	4	1	2	1	4	2	6	6	3	2	4	7	2
2 Outputs																										
2.1 Review form	•	•		•	•	•	•	•						•				•	•	•			•	•	•	•
2.2 Review report	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•	•		•	•	•	•
2.3 Code review report												•										•				
3 Activities																										
3.1 Planning	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.2 Review	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
3.3 Update during meeting	•																									
3.4 Update after meeting		•	•		•	•	•	•	•				•	•				•	•	•			•	•	•	•
3.5 Closure	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4 Roles																										
4.1 Project Manager	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.2 QA Expert	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.3 CM Specialist	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
4.4 Customer				•																						
5 Tools and Techniques																										
NOT IDENTIFIABLE																										
Process Cluster	E	A	I	H	A	A	A	A	I	B	B	D	I	A	B	B	B	A	A	A	B	B	A	A	A	A
Process Sub-cluster	5	1	1	1	4	7	2	4	1	2	2	3	1	4	1	2	1	4	2	6	6	3	2	4	7	2

Figure D.6 Process Similarity Matrix for Review Process

PE100	PE101	PE102	PE103	PE104	PE105	PE106	PE107	PE108	PE109	PE110	PE111	PE112	PE113	PE114	PE115	PE116	PE117	PE118	PE119	PE120	PE121	PE122	PE123	PE124	PE125	PE126	PE127	PE128	PE129	PE130	PE131	PE132	PE133	PE134	PE135	PE136	PE137	PE138	PE139	PE140	PE141	PE142	PE143	PE144
SRS	STP	STD	SDD	SPS	SVD	YTR	SRS	SDD	STP	STD	SPS	SVD	YTR	STP	STD	SDD	SPS	SRS	SVD	SRS	SDD	STD-KMM	YTI	YTI	QTPR	QTP	QTPR	YUO	YSI	STR	YUO	YSI	STR	YUO	YSI	YTR	YKK	YKK	YKK	STP	STD	SPS	STR	QTP
4	1	2	2	4	7	6	4	2	1	2	4	7	6	1	2	2	4	4	7	4	2	2	5	5	5	1	5	4	7	6	4	7	6	4	7	6	8	8	8	1	2	4	6	1
•	•	•	•	•		•	•	•	•	•	•	•											•	•	•	•	•						•				•	•	•				•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
																									•																			•
•	•	•	•	•		•	•	•	•	•	•	•											•	•		•	•						•				•	•	•				•	
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
																									•	•	•											•						
A	A	A	A	A	B	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	A	A	G	C	C	B	B	B	B	B	A	B	B	B	A	C	A	B	B	B	A	E
4	1	2	2	4	7	6	4	2	1	2	4	7	6	1	2	2	4	4	7	4	2	2	5	5	5	1	5	4	7	6	4	7	6	4	7	6	8	8	8	1	2	4	6	1

PE145	PE146	PE147	PE148	PE149	PE150	PE151	PE152	PE153	PE154	PE155	PE156	PE157	PE158	PE159	PE160	PE161	PE162	PE163	PE164	PE165	PE166	PE167	PE168	PE169	PE170	PE171	PE172	PE173	PE174	PE175	PE176	PE177	PE178	PE179	PE180	PE181	PE182	PE183	PE184	PE185	PE186	PE187	PE188	PE189	PE190	PE191	PE192
QTPR	STR	QTP	QTPR	QTP	QTPR	SSS	SSDD	QAP	SSS	SSDD	STR	STR	PYP	SEMP	OCD	OCD	SDP	ASYP	GST	SEMP	YGD	AKD4	QAP	CMP	CMP	KCD	YGD	YTP	QTPR	SSS	QTPR	SSS	PYP	YKP Bimal exi	YGP	YKP	KYP	YTP	SSDD	YTTD	YTTD	YUO	YSI	YTR	YKK	YKP Bimal exi	YKP Bimal exi
5	6	1	5	1	5	4	3	1	4	2	6	6	1	1	4	4	1	1	8	1	4	2	1	1	1	2	4	1	5	2	5	2	1	3	1	1	1	1	2	5	5	4	7	6	8	3	3
•						•			•	•	•		•	•	•		•	•	•		•	•	•	•		•		•				•	•		•		•		•	•	•	•	•	•	•		
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	•		
																																		•												•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•																	•																•		•		•			•	•	•	•	•	•		
	•					•		•	•	•	•		•	•	•			•	•		•	•	•	•		•		•			•	•							•								
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
		•	•																													•							•								
E	I	F	F	B	B	A	B	I	A	A	A	B	A	A	A	B	E	A	A	B	A	A	A	A	B	A	B	A	B	B	I	C	E	D	E	B	E	B	C	E	E	E	E	E	E	D	D
5	6	1	5	1	5	4	3	1	4	2	6	6	1	1	4	4	1	1	8	1	4	2	1	1	1	2	4	1	5	2	5	2	1	3	1	1	1	1	2	5	5	4	7	6	8	3	3

Figure D.6 Process Similarity Matrix for Review Process (continued)

Metric Usability Questionnaires

Metric Name: Opening date			**Assessed On:** 11/10/2005	
Conceptual Definition: Opening date of a review report			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**0,00**
	Q1	Which entity does the metric measure?	Review Report	
	Q2	Which attribute of the entity does the metric measure?	Opening date	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Interval	0
	Q4	What is the unit of the metric data?	Date	
	Q5	What is the type of the metric data? (integer, real, etc.)	Date	
	Q6	What is the range of the metric data?	Infinite	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	2	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At start	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,50**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for QMS)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	No	0
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	No	0
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q31	Is metric data integrable at project level?	Yes	
	Q32	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,00**

Figure D.7 Metric Usability Questionnaire for “Opening Date” Base Metric of Review Process

Metric Name: Closure date			**Assessed On:** 11/10/2005	
Conceptual Definition: Closure date of a review report			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**0,00**
	Q1	Which entity does the metric measure?	Review Report	
	Q2	Which attribute of the entity does the metric measure?	Closure date	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Interval	0
	Q4	What is the unit of the metric data?	Date	
	Q5	What is the type of the metric data? (integer, real, etc.)	Date	
	Q6	What is the range of the metric data?	Infinite	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	3	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the updates	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,50**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for QMS)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	No	0
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	No	0
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q31	Is metric data integrable at project level?	Yes	
	Q32	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,00**

Figure D.8 Metric Usability Questionnaire for “Closure Date” Base Metric of Review Process

Metric Name: Number of detected nonconformances			**Assessed On:** 11/10/2005	
Conceptual Definition: No.of nonconformances detected by reviewers (prior to review)			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	Which entity does the metric measure?	Product under review	
	Q2	Which attribute of the entity does the metric measure?	Quality	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Absolute	1
	Q4	What is the unit of the metric data?	Number	
	Q5	What is the type of the metric data? (integer, real, etc.)	Integer	
	Q6	What is the range of the metric data?	[0, infinity]	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	4	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the review meeting	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,50**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for QMS)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	No	0
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	No	0
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q31	Is metric data integrable at project level?	Yes	
	Q32	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,50**

Figure D.9 Metric Usability Questionnaire for "Number of Detected Nonconformances" Base Metric of Review Process

Metric Name: Number of accepted nonconformances			**Assessed On:** 11/10/2005	
Conceptual Definition: No.of nonconformances accepted in a review			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Definition				**1,00**
	Q1	Which entity does the metric measure?	Product under review	
	Q2	Which attribute of the entity does the metric measure?	Quality	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Absolute	1
	Q4	What is the unit of the metric data?	Number	
	Q5	What is the type of the metric data? (integer, real, etc.)	Integer	
	Q6	What is the range of the metric data?	[0, no.of accepted nonconformances]	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	4	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the review meeting	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,75**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for nonconformance resolution)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	Yes (as total nonconformances by projects)	1
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	Yes	1
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Is metric data normalizable per product type basis?	Yes	
	Q31	Is metric data normalizable per process phase basis?	No	
	Q32	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q33	Is metric data integrable at project level?	Yes	
	Q34	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,75**

Figure D.10 Metric Usability Questionnaire for "Number of Accepted Nonconformances" Base Metric of Review Process

Metric Name: Review effort			**Assessed On:** 11/10/2005	
Conceptual Definition: Total effort spent for product review			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	Which entity does the metric measure?	Human resource	
	Q2	Which attribute of the entity does the metric measure?	Effort	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q4	What is the unit of the metric data?	Man-hour	
	Q5	What is the type of the metric data? (integer, real, etc.)	Real	
	Q6	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	63	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	Yes (1-58)	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the review meeting	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,75**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for QMS)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	Yes (as total review effort by projects)	1
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	Yes	1
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q31	Is metric data integrable at project level?	Yes	
	Q32	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,75**

Figure D.11 Metric Usability Questionnaire for “Review Effort” Base Metric of Review Process

Metric Name: Nonconformance resolution effort			**Assessed On:** 11/10/2005	
Conceptual Definition: Total effort spent for nonconformance resolutic			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	Which entity does the metric measure?	Human resource	
	Q2	Which attribute of the entity does the metric measure?	Effort	
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q4	What is the unit of the metric data?	Man-hour	
	Q5	What is the type of the metric data? (integer, real, etc.)	Real	
	Q6	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q7	Is metric data existent?	Yes	1
	Q8	What is the amount of overall observations?	199	
	Q9	What is the amount of missing data points?	43	
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q11	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the updates	
	Q12	Is all metric data recorded at the same place in the process? (at start, middle, end, later, etc.)	Yes	1
	Q13	Who is responsible for recording metric data?	Quality Assurance Representative	
	Q14	Is all metric data recorded by the responsible body?	Yes	1
	Q15	How is metric data recorded? (on a form, report, tool, etc.)	On a report	
	Q16	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes	1
	Q17	Where is metric data stored? (in a file, database, etc.)	On review records	
	Q18	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**0,75**
	Q19	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q20	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q21	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly, etc.)	Asynchronous	
	Q22	Are the frequencies for data generation, recording, and storing different?	No	1
	Q23	Is metric data recorded precisely?	Yes	1
	Q24	Is metric data collected for a specific purpose?	Yes (for QMS)	1
	Q25	Is the purpose of metric data collection known by process performers?	Yes	1
	Q26	Is metric data analyzed and reported?	Yes (as total resolution effort by projects)	1
	Q27	Is metric data analysis results communicated to process performers?	No	0
	Q28	Is metric data analysis results communicated to management?	Yes	1
	Q29	Is metric data analysis results used as a basis for decision making?	No	0
Data Normalizability				
	Q30	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q31	Is metric data integrable at project level?	Yes	
	Q32	Is metric data integrable at organization level?	Yes	
			Total Points (Metric Usability Index)	**0,75**

Figure D.12 Metric Usability Questionnaire for “Nonconformance Resolution Effort” Base Metric of Review Process

Metric Name: Noncoformance detection efficiency			**Assessed On:** 11/10/2005	
Conceptual Definition: Average effort to detect a nonconformance			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	What is the the metric formula? (please refer to related base metrics)	Number of accepted nonconformances / Review effort	
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q3	What is the unit of the metric data?	Number of nonconformances per man-hour	
	Q4	What is the type of the metric data? (integer, real, etc.)	Real	
	Q5	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q6	Is metric data existent?	Yes	1
	Q7	What is the amount of overall observations?	199	
	Q8	What is the amount of missing data points?	63	
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	Yes (1-58)	
Data Verifiability				**1,00**
	Q10	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula	
	Q11	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes	1
	Q12	Is all metric data calculated according to metric formula?	Yes	1
	Q13	Where is metric data stored? (in a file, database, etc.)	In excel sheet	
	Q14	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**N/A**
	Q15	Is metric data stored precisely?	Yes	1
	Q16	Is metric data stored for a specific purpose?	Yes	1
	Q17	Is the purpose of metric data storage known by process performers?	N/A	
	Q18	Is metric data analyzed and reported?	N/A	
	Q19	Is metric data analysis results communicated to process performers?	N/A	
	Q20	Is metric data analysis results communicated to management?	N/A	
	Q21	Is metric data analysis results used as a basis for decision making?	N/A	
Data Normalizability				
	Q22	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q23	Is metric data integrable at project level?	Yes	
	Q24	Is metric data integrable at organization level?	Yes	
			Total Points	**1,00**
			MUI for no.of accepted nonconformances	**0,75**
			MUI for review effort	**0,75**
			Metric Usability Index	**0,75**

Figure D.13 Metric Usability Questionnaire for “Nonconformance Detection Efficiency” Derived Metric of Review Process

Metric Name: Noncoformance resolution efficiency			**Assessed On:** 11/10/2005	
Conceptual Definition: Average effort to remove a nonconformance			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	What is the the metric formula? (please refer to related base metrics)	Number of accepted nonconformances / Nonconformance resolution effort	
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q3	What is the unit of the metric data?	Number of nonconformances per man-hour	
	Q4	What is the type of the metric data? (integer, real, etc.)	Real	
	Q5	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q6	Is metric data existent?	Yes	1
	Q7	What is the amount of overall observations?	199	
	Q8	What is the amount of missing data points?	110	
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q10	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula	
	Q11	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes	1
	Q12	Is all metric data calculated according to metric formula?	Yes	1
	Q13	Where is metric data stored? (in a file, database, etc.)	In excel sheet	
	Q14	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**N/A**
	Q15	Is metric data stored precisely?	Yes	1
	Q16	Is metric data stored for a specific purpose?	Yes	1
	Q17	Is the purpose of metric data storage known by process performers?	N/A	
	Q18	Is metric data analyzed and reported?	N/A	
	Q19	Is metric data analysis results communicated to process performers?	N/A	
	Q20	Is metric data analysis results communicated to management?	N/A	
	Q21	Is metric data analysis results used as a basis for decision making?	N/A	
Data Normalizability				
	Q22	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q23	Is metric data integrable at project level?	Yes	
	Q24	Is metric data integrable at organization level?	Yes	
			Total Points	**1,00**
			MUI for no.of accepted nonconformances	**0,75**
			MUI for nonconformance resolution effort	**0,75**
			Metric Usability Index	**0,75**

Figure D.14 Metric Usability Questionnaire for "Nonconformance Resolution Efficiency" Derived Metric of Review Process

Metric Name: Open period			**Assessed On:** 11/10/2005	
Conceptual Definition: Duration that a review report stays open			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	What is the the metric formula? (please refer to related base metrics)	Closure date - Opening date	
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Absolute	1
	Q3	What is the unit of the metric data?	Day	
	Q4	What is the type of the metric data? (integer, real, etc.)	Integer	
	Q5	What is the range of the metric data?	[0, infinity]	
Data Existence				**1,00**
	Q6	Is metric data existent?	Yes	1
	Q7	What is the amount of overall observations?	199	
	Q8	What is the amount of missing data points?	3	
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q10	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula	
	Q11	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes	1
	Q12	Is all metric data calculated according to metric formula?	Yes	1
	Q13	Where is metric data stored? (in a file, database, etc.)	In excel sheet	
	Q14	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**N/A**
	Q15	Is metric data stored precisely?	Yes	1
	Q16	Is metric data stored for a specific purpose?	Yes	1
	Q17	Is the purpose of metric data storage known by process performers?	N/A	
	Q18	Is metric data analyzed and reported?	N/A	
	Q19	Is metric data analysis results communicated to process performers?	N/A	
	Q20	Is metric data analysis results communicated to management?	N/A	
	Q21	Is metric data analysis results used as a basis for decision making?	N/A	
Data Normalizability				
	Q22	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes. Number of accepted detected nonconformances.	
Data Integrability				
	Q23	Is metric data integrable at project level?	Yes	
	Q24	Is metric data integrable at organization level?	Yes	
			Total Points	**1,00**
			MUI for opening date	0,50
			MUI for closure date	0,50
			Metric Usability Index	**0,50**

Figure D.15 Metric Usability Questionnaire for "Open Period" Derived Metric of Review Process

Metric Name: Open period with respect to nonconformances			**Assessed On:** 11/10/2005	
Conceptual Definition: Open period normalized by no.of accepted nonconformances			**Assessed By:** Ayça Tarhan	
Attributes			**Answers**	**Points**
Metric Identity				**1,00**
	Q1	What is the the metric formula? (please refer to related base metrics)	Open period / Number of accepted nonconformances	
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio	1
	Q3	What is the unit of the metric data?	Days per nonconformance	
	Q4	What is the type of the metric data? (integer, real, etc.)	Real	
	Q5	What is the range of the metric data?	[0.0, infinity]	
Data Existence				**1,00**
	Q6	Is metric data existent?	Yes	1
	Q7	What is the amount of overall observations?	199	
	Q8	What is the amount of missing data points?	90	
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing periods)	No	
Data Verifiability				**1,00**
	Q10	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula	
	Q11	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes	1
	Q12	Is all metric data calculated according to metric formula?	Yes	1
	Q13	Where is metric data stored? (in a file, database, etc.)	In excel sheet	
	Q14	Is all metric data stored in the same place? (in a file, database, etc.)	Yes	1
Data Dependability				**N/A**
	Q15	Is metric data stored precisely?	Yes	1
	Q16	Is metric data stored for a specific purpose?	Yes	1
	Q17	Is the purpose of metric data storage known by process performers?	N/A	
	Q18	Is metric data analyzed and reported?	N/A	
	Q19	Is metric data analysis results communicated to process performers?	N/A	
	Q20	Is metric data analysis results communicated to management?	N/A	
	Q21	Is metric data analysis results used as a basis for decision making?	N/A	
Data Normalizability				
	Q22	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No	
Data Integrability				
	Q23	Is metric data integrable at project level?	Yes	
	Q24	Is metric data integrable at organization level?	Yes	
			Total Points	**1,00**
			MUI for opening date	0,50
			MUI for closure date	0,50
			MUI for no. of accepted nonconformances	0,75
			Metric Usability Index	**0,58**

Figure D.16 Metric Usability Questionnaire for "Open Period with respect to Nonconformances" Derived Metric of Review Process

Process Execution Questionnaires

Process Name: Review		Recorded On: 03.11.2005	
Process Execution No: 51		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	Yes	Project schedule was very tight, and review time was limited.
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure D.17 Process Execution Questionnaire for Review PE # 51

Process Name: Review		Recorded On: 03.11.2005	
Process Execution No: 59, 62, 63, 64		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	Yes	Project schedule was very tight, and review time was limited.
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure D.18 Process Execution Questionnaire for Review PEs # 59, 62, 63, 64

Process Name: Review		Recorded On: 03.11.2005	
Process Execution No: 70		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	Product under review was qualification test procedures, and awaited related updates in analysis and design documents.		

Figure D.19 Process Execution Questionnaire for Review PE # 70

Process Name: Review		Recorded On: 03.11.2005	
Process Execution No: 98		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	Yes	Computer based review
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	Yes	Review was performed primarily for consistency checking (verification) - not like standard document review
OTHER FACTORS (Please list if any)			

Figure D.20 Process Execution Questionnaire for Review PE # 98

Process Name: Review		Recorded On: 03.11.2005	
Process Execution No: 156		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	Product under review was system test report and nonconformances are mostly syntactical in STR (it takes less time to rosolve).		

Figure D.21 Process Execution Questionnaire for Review PE # 156

Process Attribute Descriptions

SPC
Process Attributes Description

Process Name:	Review	Described On:	12/11/2005
Process Cluster:	M	Described By:	AT

1. **Inputs:** Please list the inputs to the process.

No	Name	Description
1	Product type under review	Can have values: 1) Project plans, 2) Design documents, 3) Analysis documents, 4) Other documents

2. **Outputs:** Please list the outputs from the process.

No	Name	Description
1	Review Form	Take care to fill it out (was missing for some executions)
2	Review Report	

3. **Activities (in sequence):** Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Planning	
2	Review	
3	Update after meeting	
4	Closure	

4. **Roles:** Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	
4	Customer	in joint reviews

5. **Tools and Techniques:** Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
	NOT RECORDED	

Figure D.22 Process Attributes Description for Review Process Cluster M

SPC[TM]
Process Attributes Description

Process Name:	Review	Described On:	12/11/2005
Process Cluster:	N	Described By:	AT

1. **Inputs:** Please list the inputs to the process.

No	Name	Description
1	Product under review	

2. **Outputs:** Please list the outputs from the process.

No	Name	Description
1	Review Form	
2	Review Report	

3. **Activities (in sequence):** Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Planning	
2	Review	
3	Update during meeting	
4	Closure	

4. **Roles:** Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Project Manager	
2	Quality Assurance Expert	
3	Configuration Management Specialist	
4	Customer	in joint reviews

5. **Tools and Techniques:** Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
	NOT RECORDED	

Figure D.23 Process Attributes Description for Review Process Cluster N

Control Charts

Nonconformance Detection Efficiency

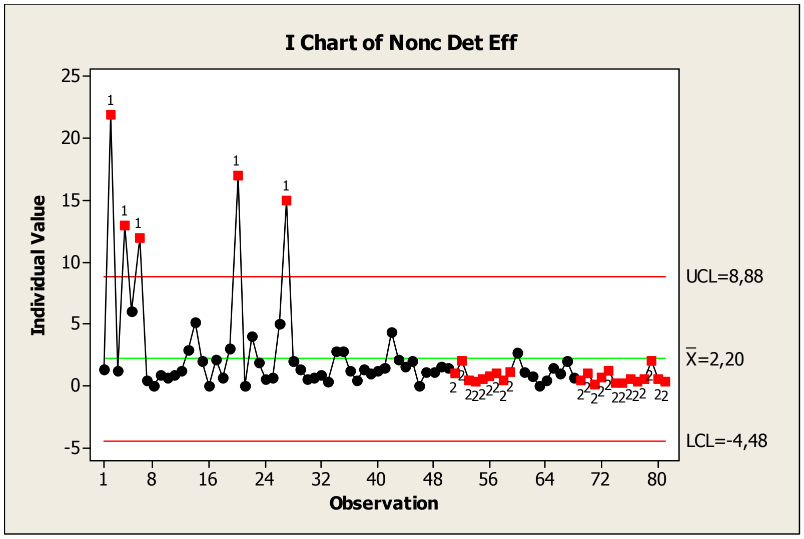


Figure D.24 Control Chart for Nonconformance Detection Efficiency of Overall Review Process

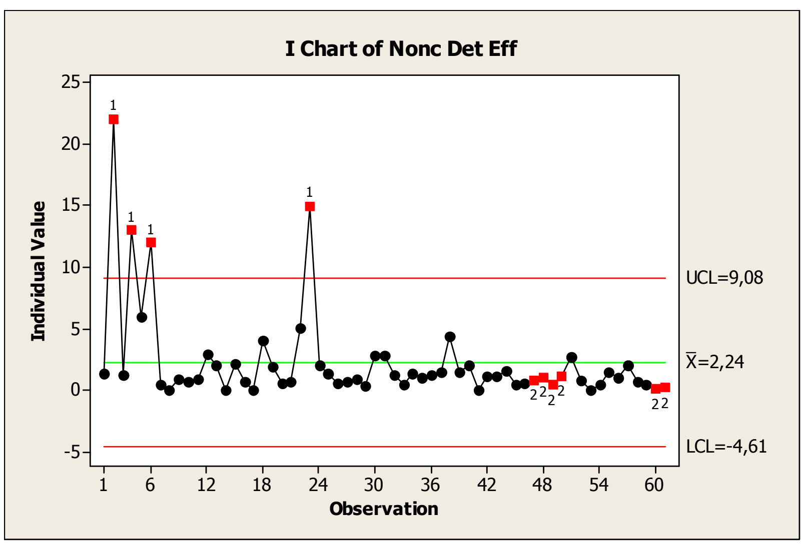


Figure D.25 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M

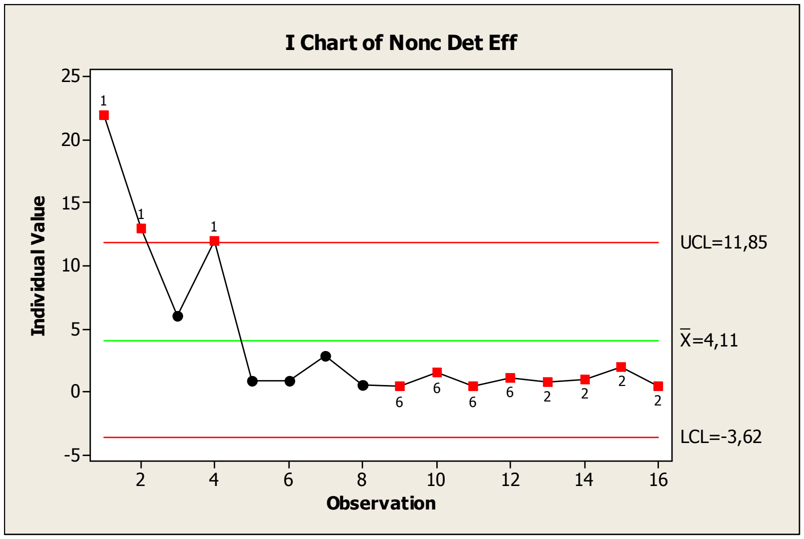


Figure D.26 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-1

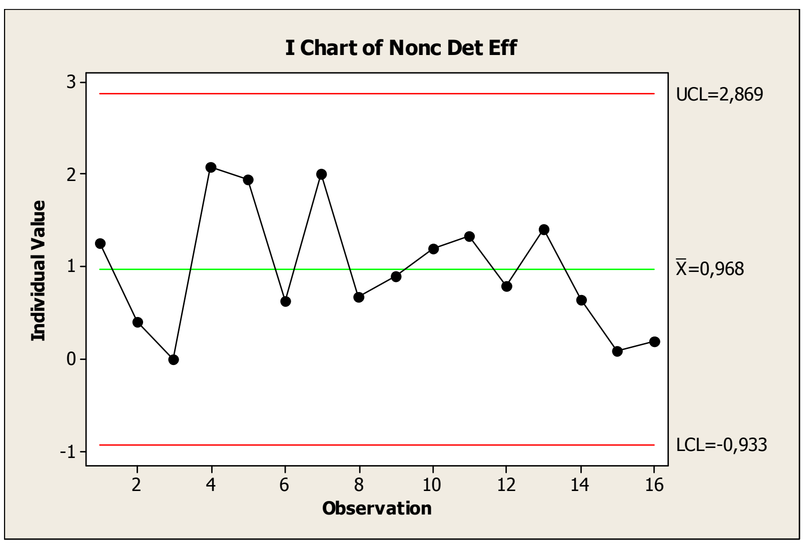


Figure D.27 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-2

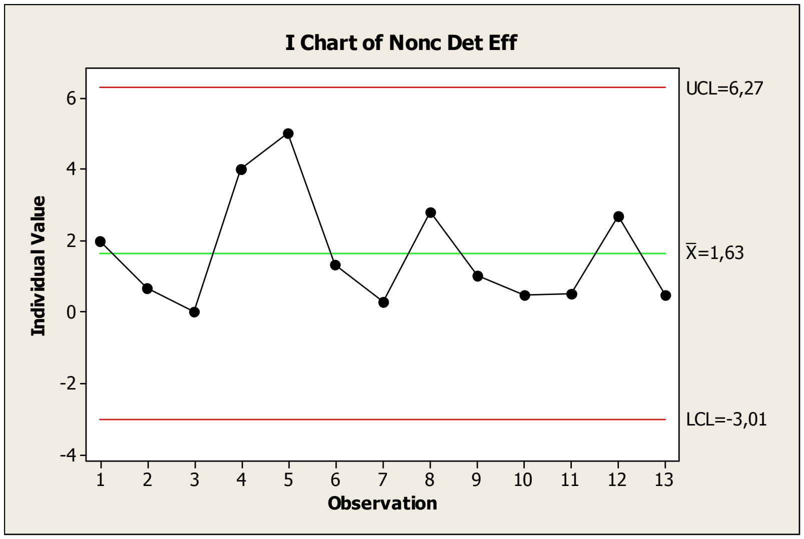


Figure D.28 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-3

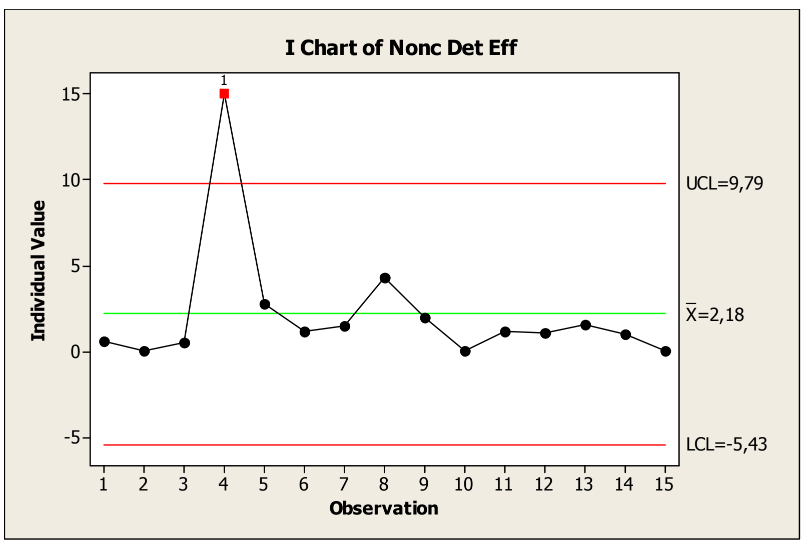


Figure D.29 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster M-4

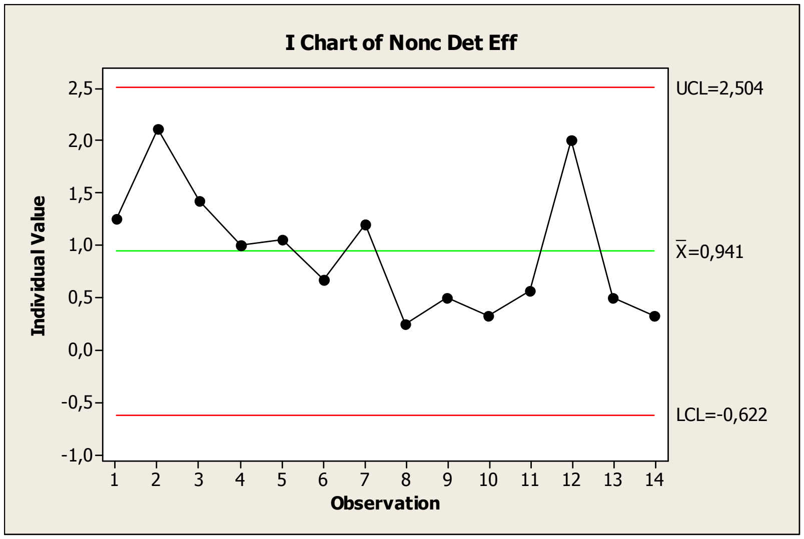


Figure D.30 Control Chart for Nonconformance Detection Efficiency of Review Process Cluster N

Nonconformance Resolution Efficiency

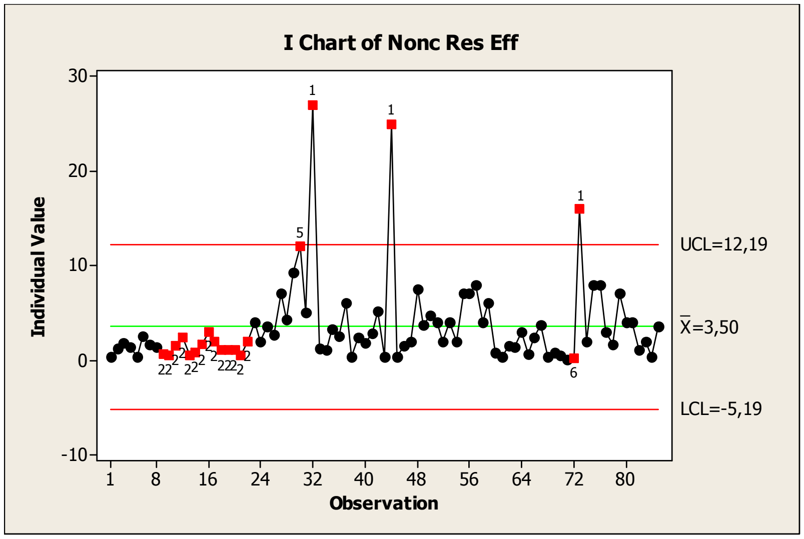


Figure D.31 Control Chart for Nonconformance Resolution Efficiency of Overall Review Process

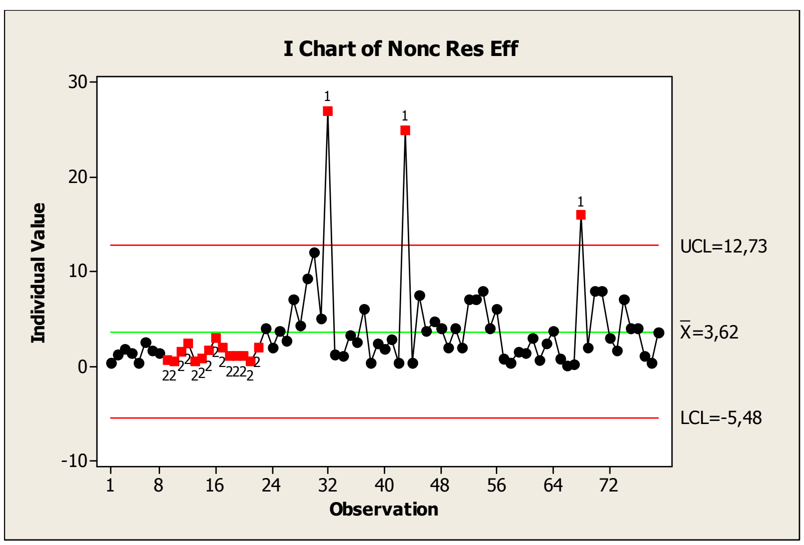


Figure D.32 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M

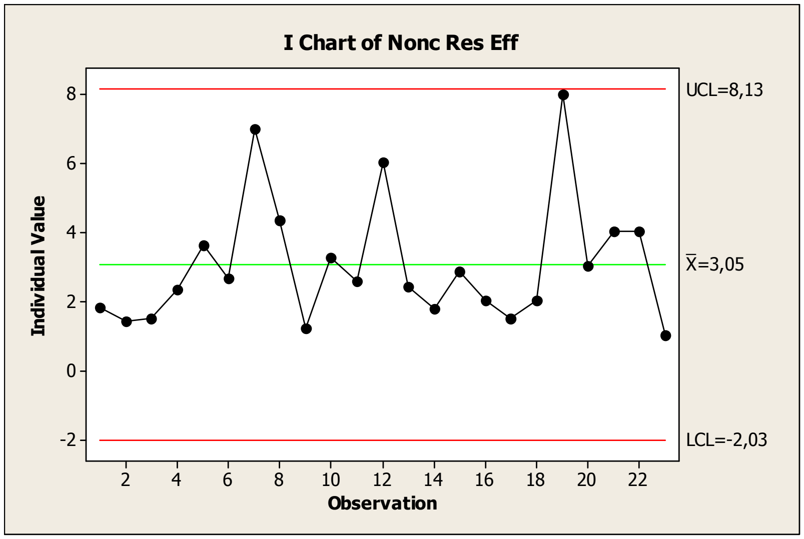


Figure D.33 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-1

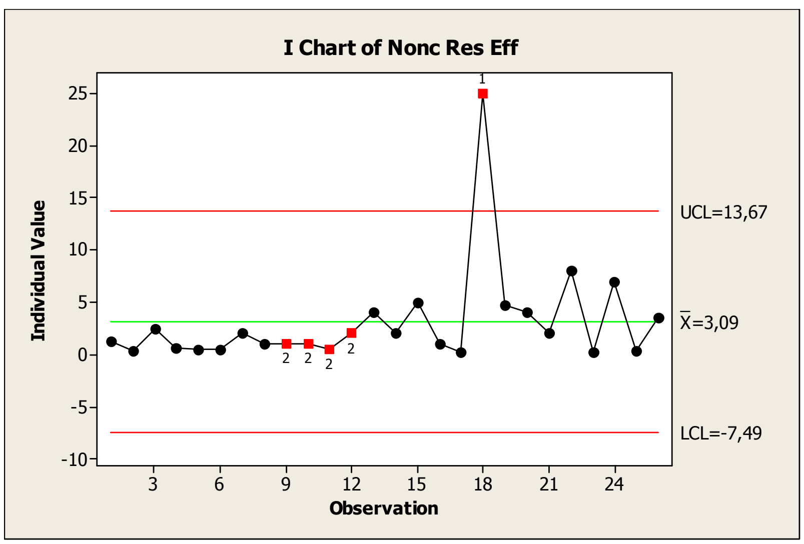


Figure D.34 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-2

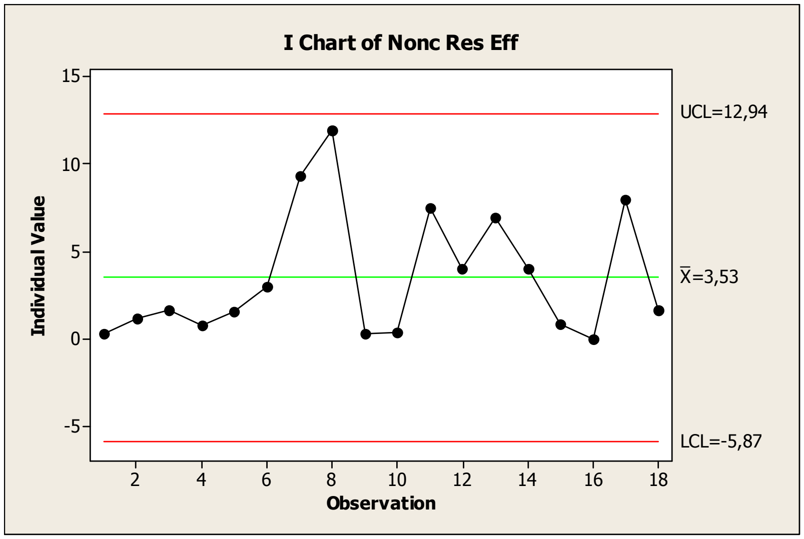


Figure D.35 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-3

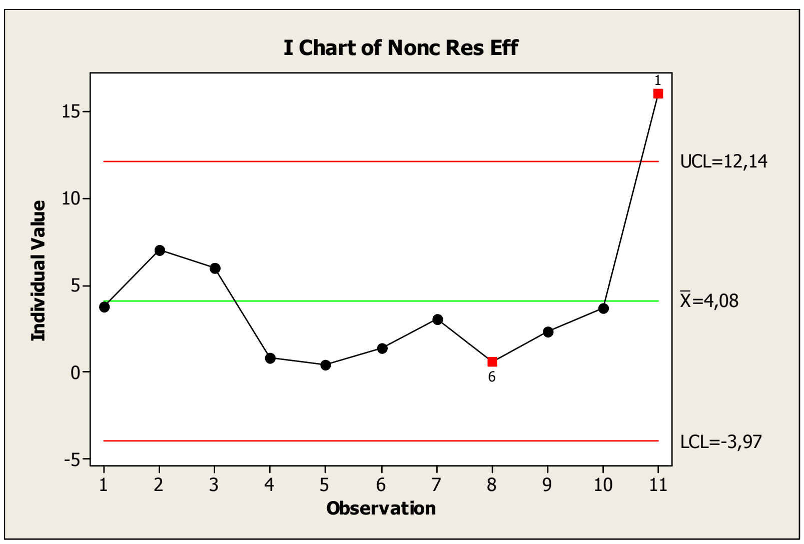


Figure D.36 Control Chart for Nonconformance Resolution Efficiency of Review Process Cluster M-4

Open Period with respect to Nonconformances

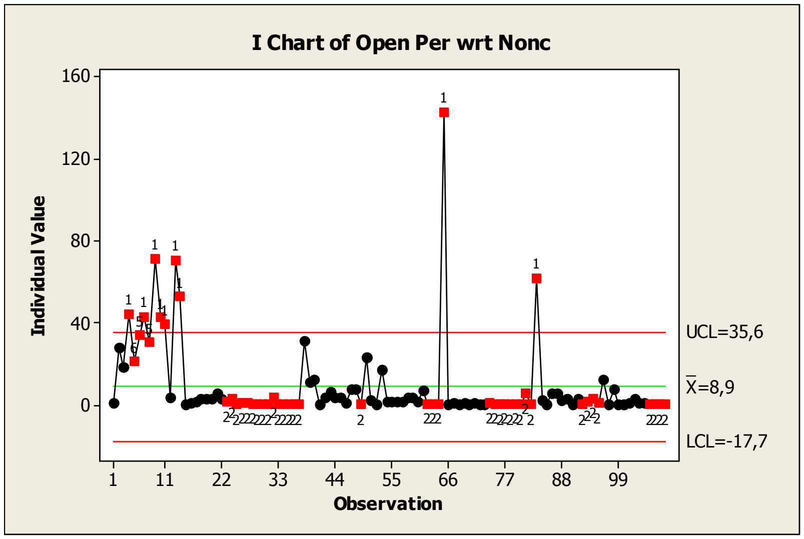


Figure D.37 Control Chart for Open Period with respect to Nonconformances of Overall Review Process

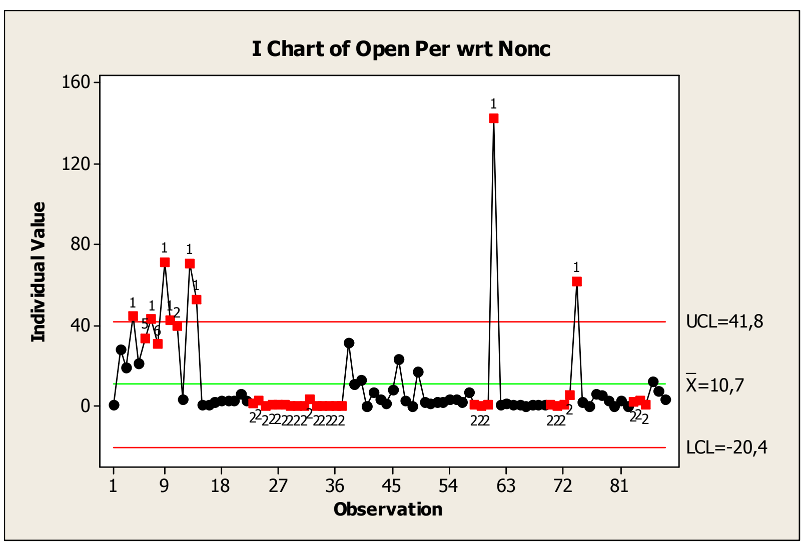


Figure D.38 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M

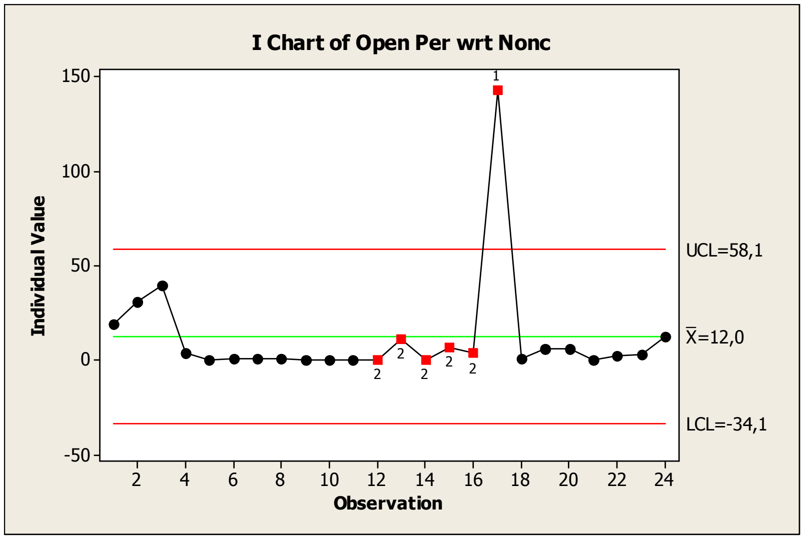


Figure D.39 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-1

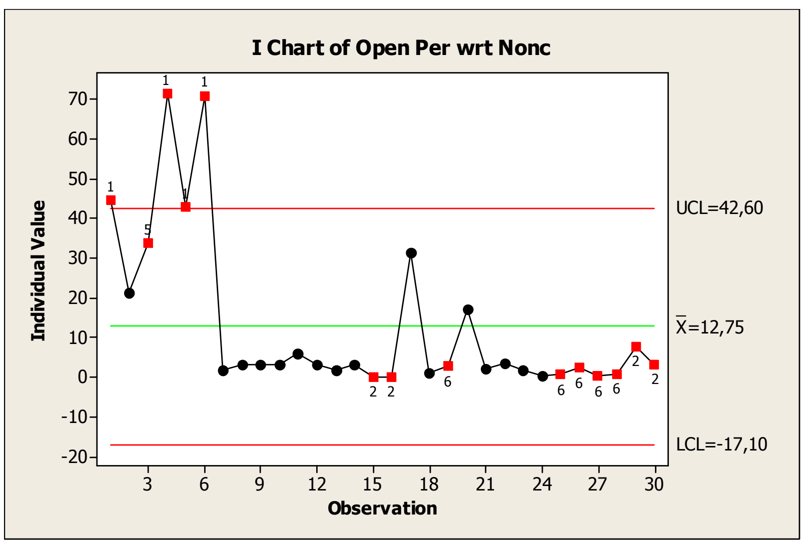


Figure D.40 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-2

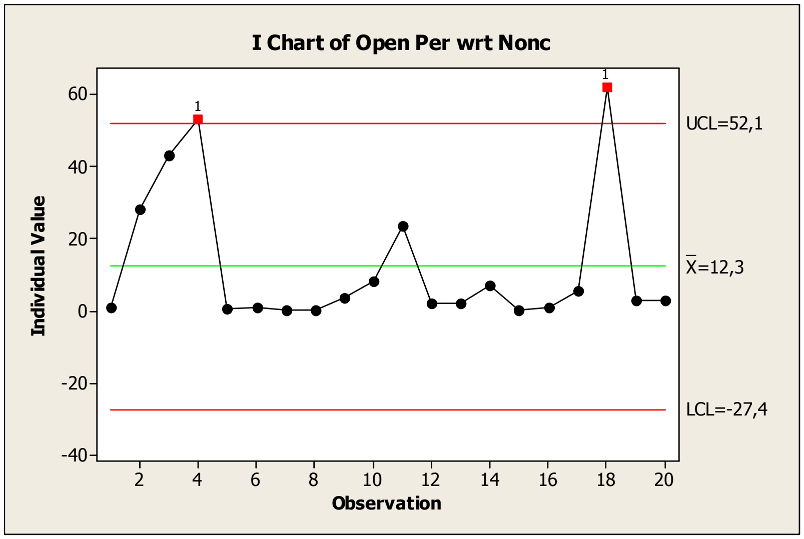


Figure D.41 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-3

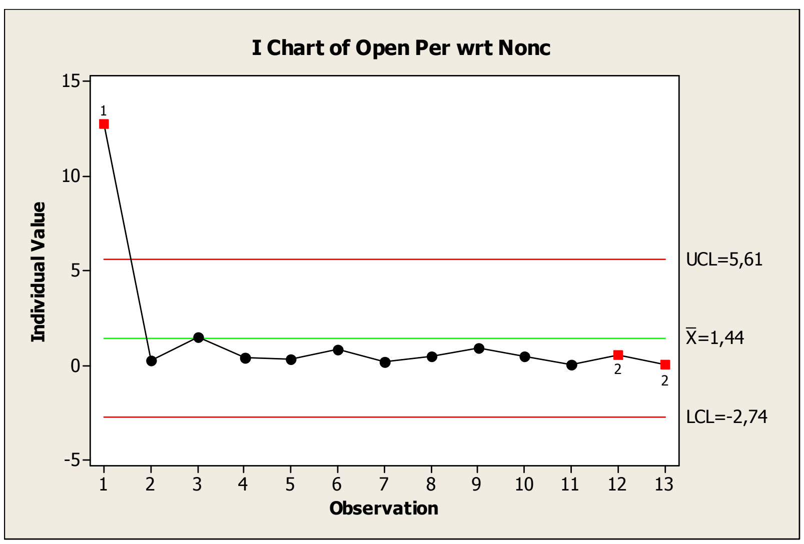


Figure D.42 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster M-4

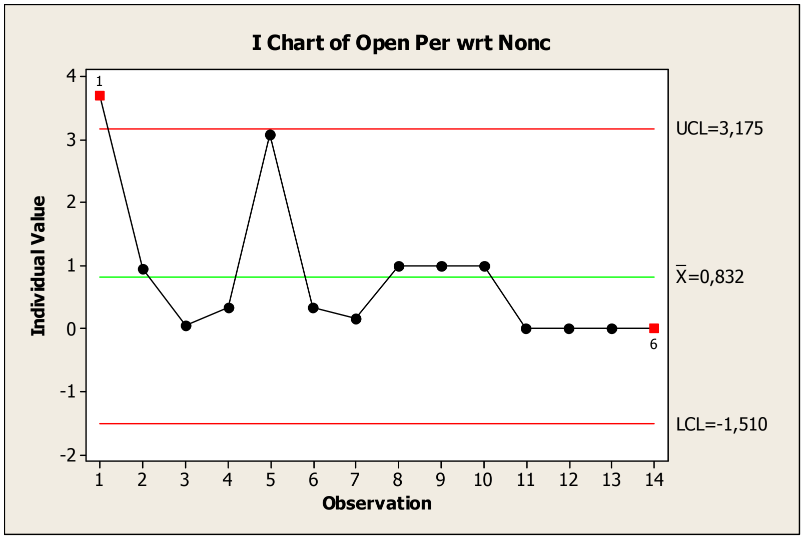


Figure D.43 Control Chart for Open Period with respect to Nonconformances of Review Process Cluster N

Open Period

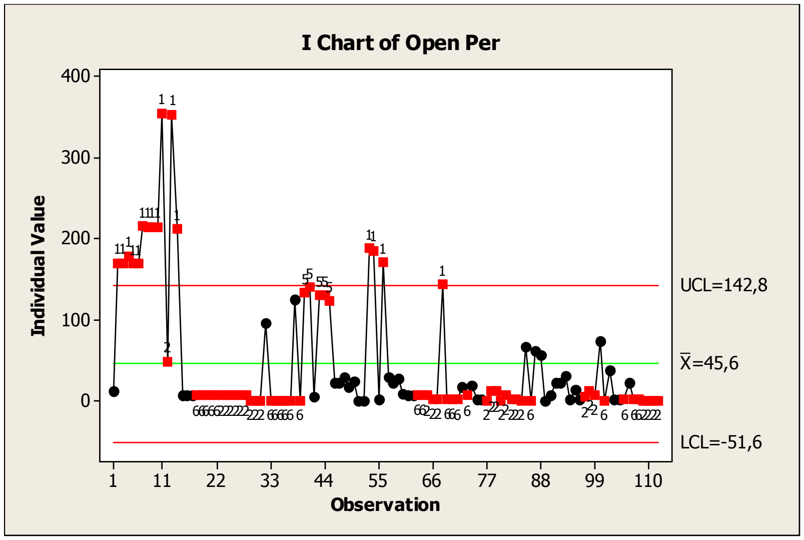


Figure D.44 Control Chart for Open Period of Overall Review Process

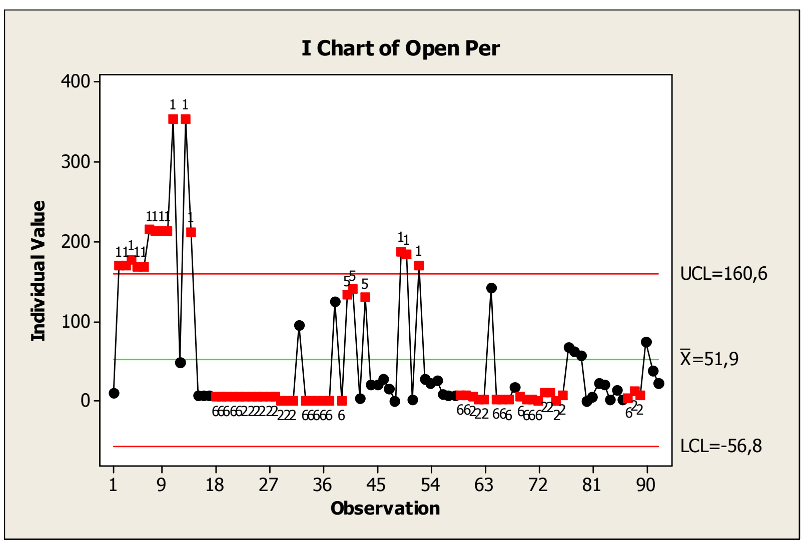


Figure D.45 Control Chart for Open Period of Review Process Cluster M

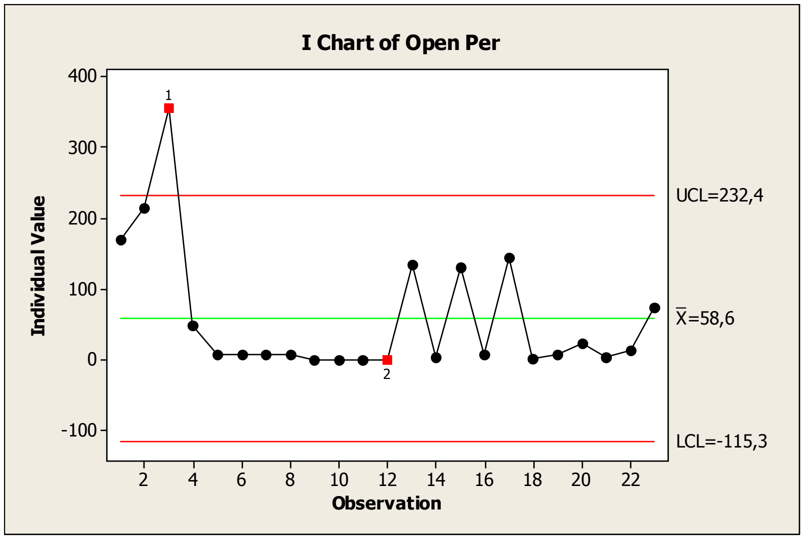


Figure D.46 Control Chart for Open Period of Review Process Cluster M-1

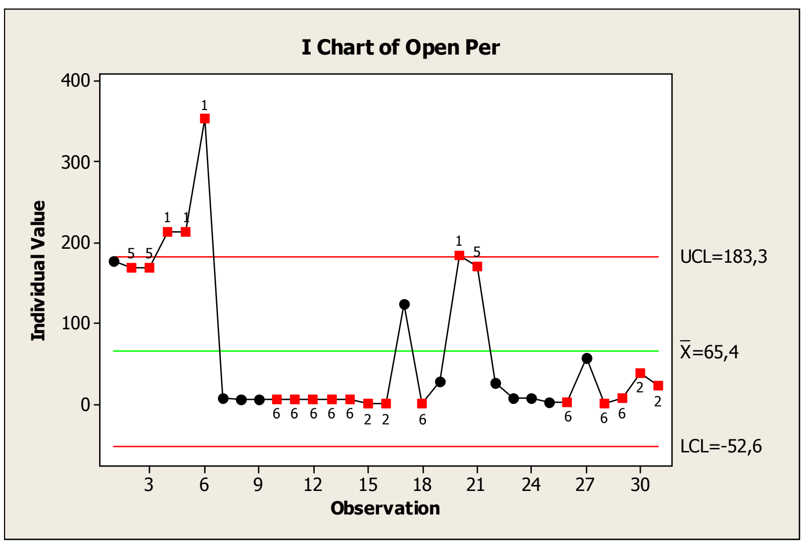


Figure D.47 Control Chart for Open Period of Review Process Cluster M-2

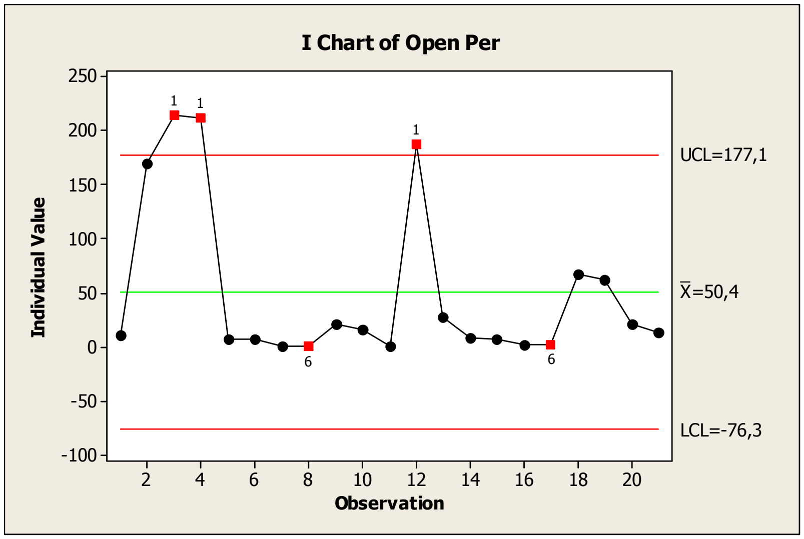


Figure D.48 Control Chart for Open Period of Review Process Cluster M-3

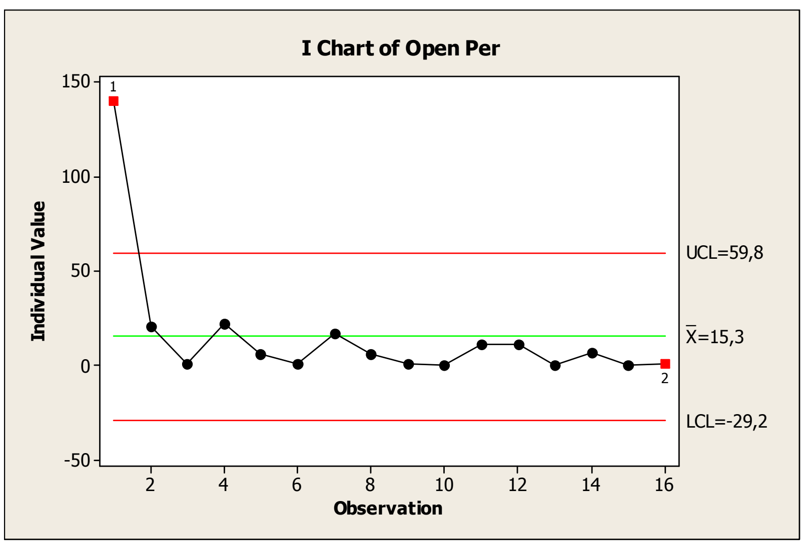


Figure D.49 Control Chart for Open Period of Review Process Cluster M-4

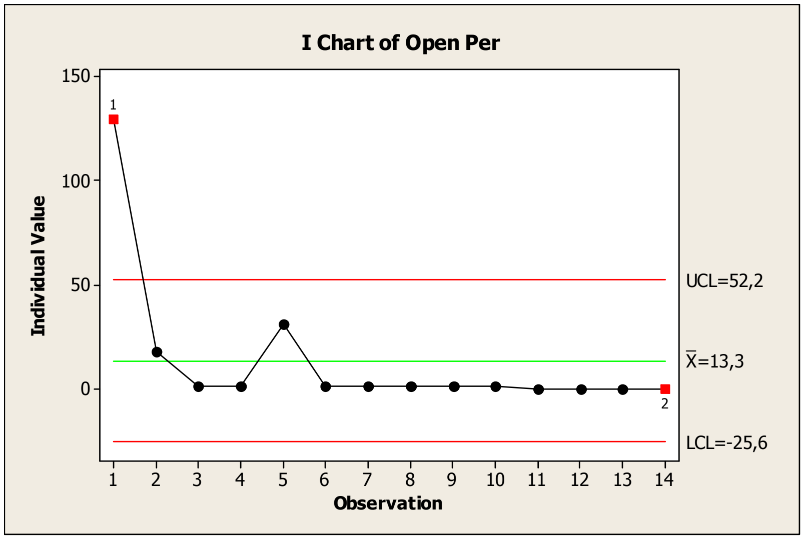


Figure D.50 Control Chart for Open Period of Review Process Cluster N

Defect Density

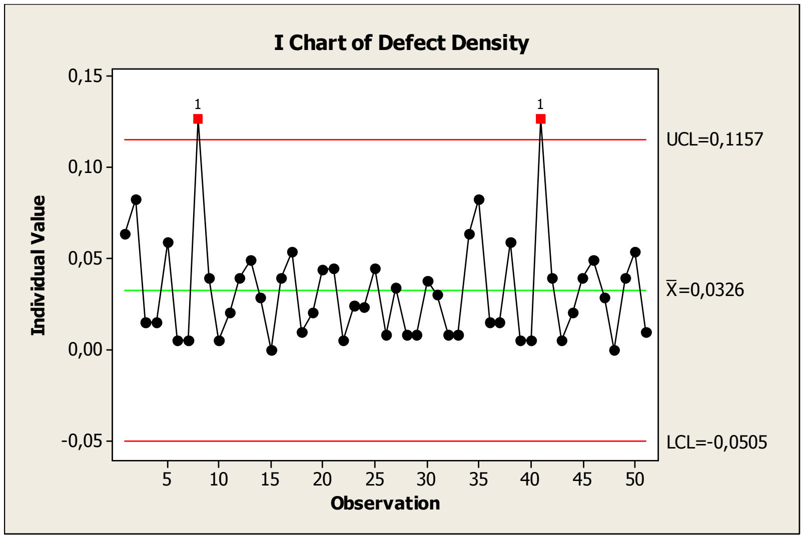


Figure D.51 Control Chart for Defect Density of Overall Review Process

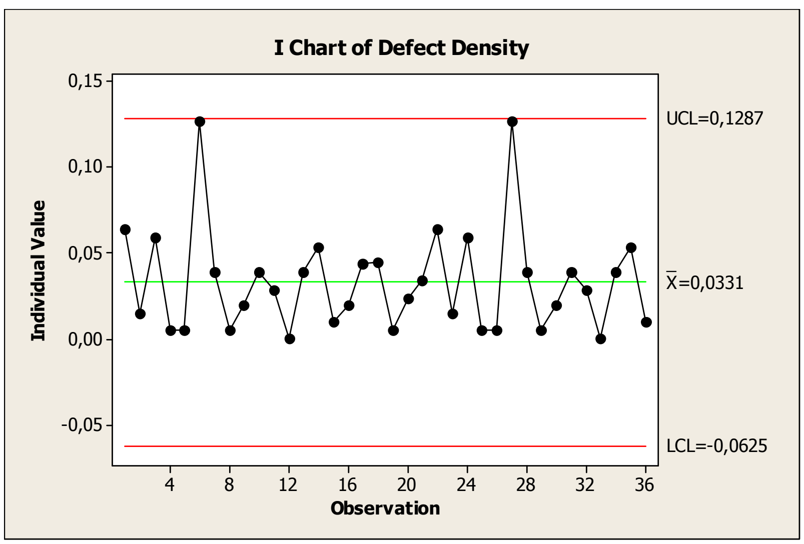


Figure D.52 Control Chart for Defect Density of Review Process Cluster M

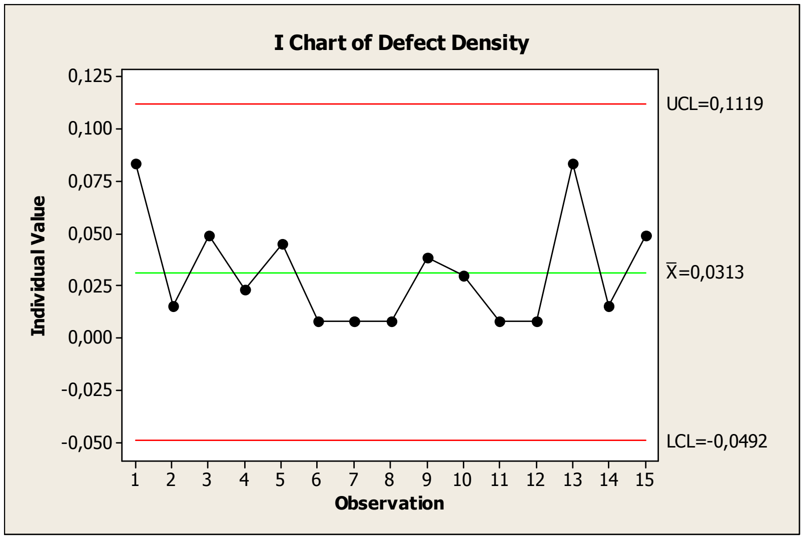


Figure D.53 Control Chart for Defect Density of Review Process Cluster N

APPENDIX E. DETAILS OF CASE STUDIES C, D, E

SPC-AM Assets

Process Execution Records

We did not fill process execution records for cases C, D, and E. Instead, we preferred drawing general process flows with the Team Leader in eEPC (Extended Event Driven Process Change) notation. Accordingly;

- The flow for test design process (case C) is given in Figure E.1.
- The flow for test procedure development process (case D) is given in Figure E.2.
- The flow for test development peer review process (case E) is given in Figure E.3.

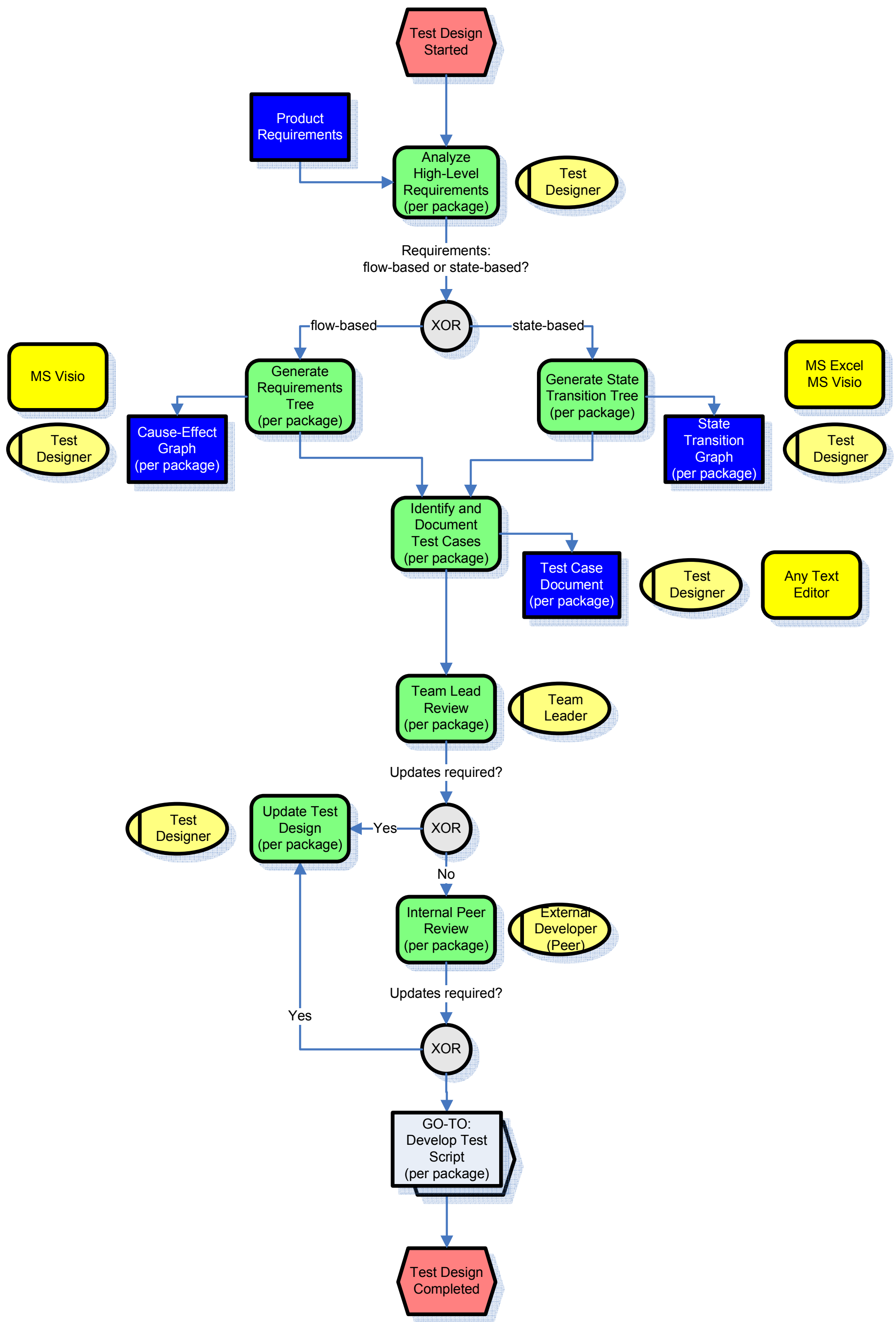


Figure E.1 Test Design Process Flow

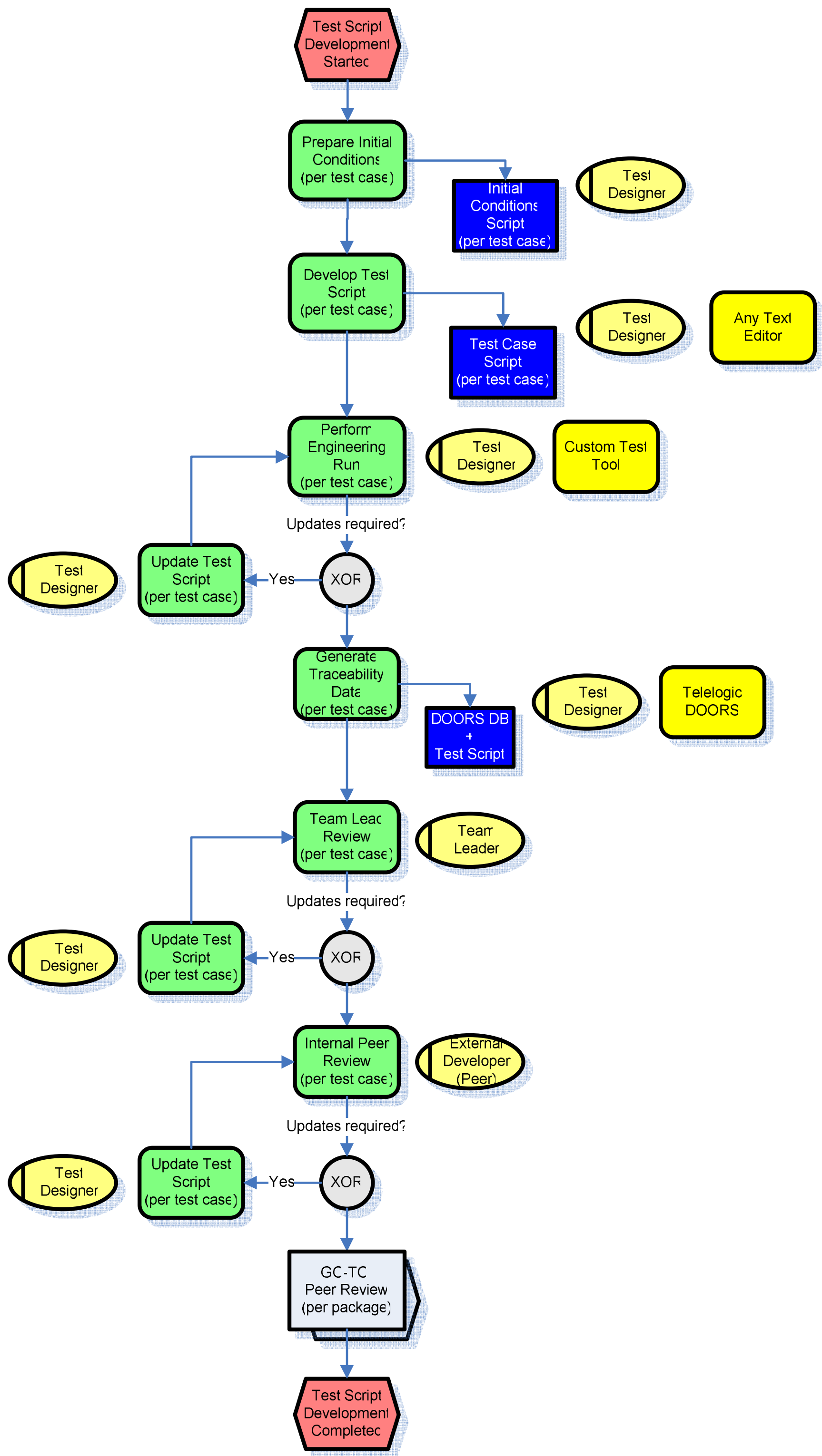


Figure E.2 Test Procedure Development Process Flow

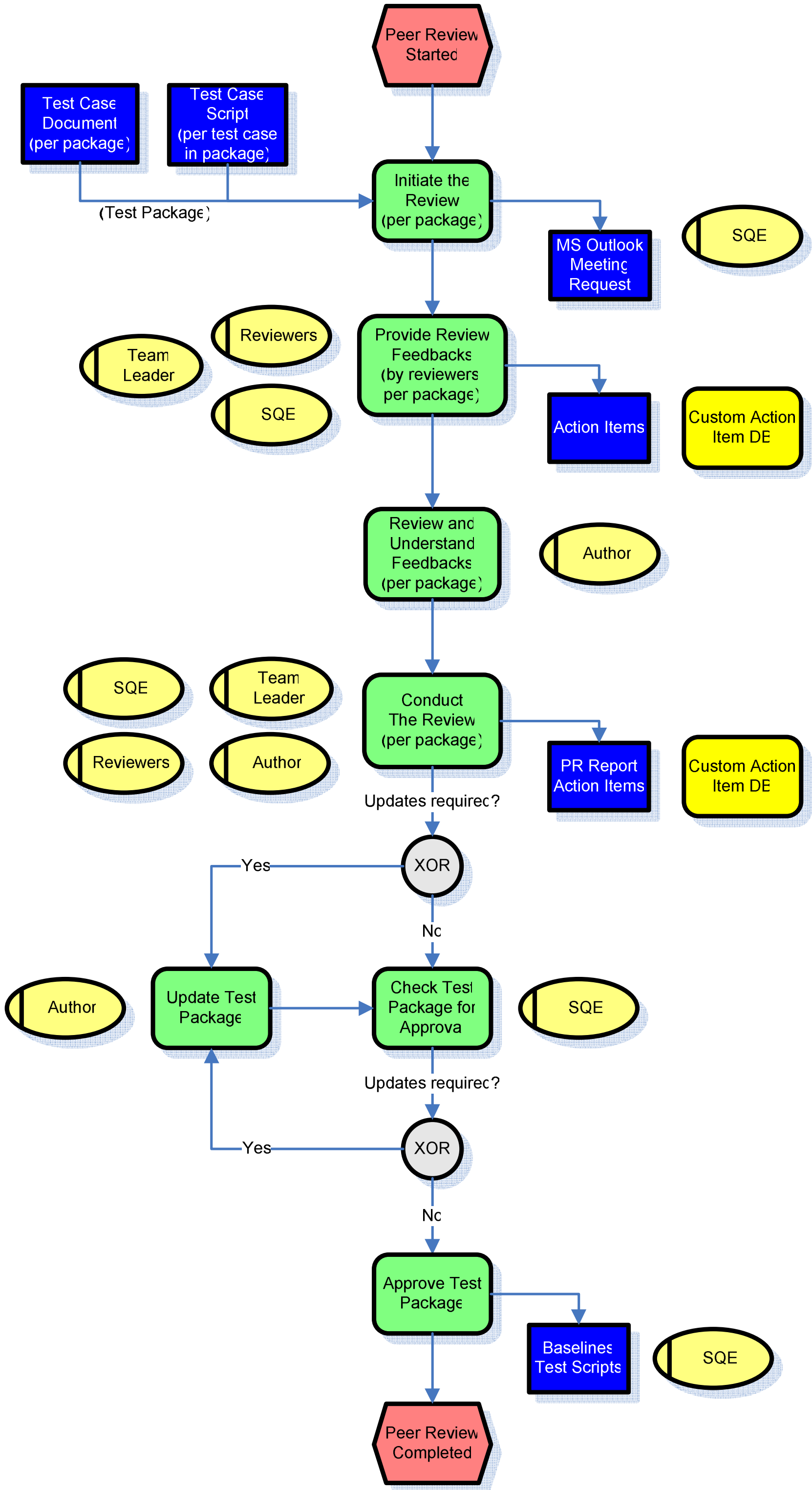


Figure E.3 Test Development Peer Review Process Flow

Process Similarity Matrices

- The matrix for test design process is given in Figure E.4.
- The matrix for test procedure development process is given in Figure E.5.
- The matrix for test development peer review process is given in Figure E.6.

PROCESS NAME: TEST DESIGN
(per package)

Please verify each process execution against process attributes. Insert an "O" into each cell if applicable (leave blank if not applicable).

Process Executions

No.	Process Attributes	PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	PE21	PE22	PE23	PE24	PE25	PE26
1	**Inputs**																										
1.1	Product Requirements	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2	**Outputs**																										
2.1	Requirements Tree										O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.2	Cause-Effect Graph										O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.3	State-Transition Graph	O	O	O	O	O	O	O	O	O																	
2.4	Test Case Document	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3	**Activities**																										
3.1	Analyze HL Requirements	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.2	Generate Test Design	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.3	Identify and Document Test Cases	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.4	Team Lead Review	O	O					O								O		O			O	O	O				
3.5	Internal Peer Review	O	O	O		O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.6	Update Test Design after IP Review	O	O	O		O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4	**Roles**																										
4.1	Test Designer	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.2	Team Leader	O	O					O								O		O			O	O	O				
4.3	External Developer (Peer)	O	O	O		O	O	O	O							O		O			O	O	O				
5	**Tools and Techniques**																										
5.1	MS Visio	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
5.2	MS Excel	O	O	O	O	O	O	O	O	O																	
5.3	Text Editor	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
	Process Cluster																										

No.	Process Attributes	PE27	PE28	PE29	PE30	PE31	PE32	PE33	PE34	PE35	PE36	PE37	PE38	PE39	PE40	PE41	PE42	PE43	PE44	PE45	PE46	PE47	PE48	PE49	PE50	PE51	PE52
1	**Inputs**																										
1.1	Product Requirements	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2	**Outputs**																										
2.1	Requirements Tree																										
2.2	Cause-Effect Graph																										
2.3	State-Transition Graph																										
2.4	Test Case Document																										
3	**Activities**																										
3.1	Analyze HL Requirements																										
3.2	Generate Test Design																										
3.3	Identify and Document Test Cases																										
3.4	Team Lead Review																										
3.5	Internal Peer Review																										
3.6	Update Test Design after IP Review																										
4	**Roles**																										
4.1	Test Designer																										
4.2	Team Leader																										
4.3	External Developer (Peer)																										
5	**Tools and Techniques**																										
5.1	MS Visio																										
5.2	MS Excel																										
5.3	Text Editor																										
	Process Cluster																										

Figure E.4 Process Similarity Matrix for Test Design Process

PROCESS NAME: TEST SCRIPT DEVELOPMENT (per test case)

Please verify each process execution against process attributes. Insert an "o" into each cell if applicable (leave blank if not applicable).

Process Executions

	Process Attributes	PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	PE21	PE22	PE23	PE24	PE25	PE26	PE27	PE28	PE29	PE30	PE31	PE32	PE33	PE34	PE35	PE36	PE37	PE38	PE39	PE40	PE41	PE42	PE43	PE44	PE45	PE46	PE47	PE48	PE49	PE50	PE51	PE52
1	**Inputs**																																																				
1.1	Test Case Document	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2	**Outputs**																																																				
2.1	Initial Conditions Script	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2.2	Test Case Script	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2.3	Traceability Data	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3	**Activities**																																																				
3.1	Prepare Initial Conditions	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.2	Develop Test Script	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.3	Perform Engineering Run	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.4	Update Test Script after Engineering Run	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.5	Generate Traceability Data	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.6	Team Lead Review	o	o	o	o	o	o	o	o	o	o		o	o	o	o	o	o	o		o	o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.7	Internal Peer Review	o	o	o	o	o	o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3.8	Update Test Script after IP Review	o	o	o	o	o	o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4	**Roles**																																																				
4.1	Test Designer	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4.2	Team Leader	o	o	o	o	o	o	o	o	o	o		o	o	o	o	o	o	o		o	o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4.3	External Developer (Peer)	o	o	o	o	o	o	o	o	o	o		o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
5	**Tools and Techniques**																																																				
5.1	Text Editor	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
5.2	Custom Test Tool	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
5.3	Telelogic DOORS	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
	Process Cluster																																																				

Figure E.5 Process Similarity Matrix for Test Procedure Development Process

PROCESS NAME: PEER REVIEW
(per package)

Please verify each process execution against process attributes. Insert an "O" into each cell if applicable (leave blank if not applicable).

Process Executions

Process Attributes	PE1	PE2	PE3	PE4	PE5	PE6	PE7	PE8	PE9	PE10	PE11	PE12	PE13	PE14	PE15	PE16	PE17	PE18	PE19	PE20	PE21	PE22	PE23	PE24	PE25	PE26
1 Inputs																										
1.1 Test Case Scripts	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
1.2 Product Requirements	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2 Outputs																										
2.1 Meeting Request	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.2 Action Items	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.3 Peer Review Report	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.4 Baseline Test Case Scripts	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3 Activities																										
3.1 Initiate the Review	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.2 Provide Review Feedbacks	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.3 Review and Understand Feedbacks	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.4 Conduct the Review	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.5 Update Test Scripts after Review	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.6 Check Test Scripts for Approval	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.7 Update Test Scripts after Approval Checking	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.8 Approve Test Scripts	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4 Roles																										
4.1 SQE	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.3 Team Leader	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.4 Reviewers	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.5 Author	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
5 Tools and Techniques																										
5.1 MS Outlook	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
5.2 Custom Action Item DB	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
Process Cluster																										

Process Attributes	PE27	PE28	PE29	PE30	PE31	PE32	PE33	PE34	PE35	PE36	PE37	PE38	PE39	PE40	PE41	PE42	PE43	PE44	PE45	PE46	PE47	PE48	PE49	PE50	PE51	PE52
1 Inputs																										
1.1 Test Case Scripts	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
1.2 Product Requirements	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2 Outputs																										
2.1 Meeting Request	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.2 Action Items	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.3 Peer Review Report	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
2.4 Baseline Test Case Scripts	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3 Activities																										
3.1 Initiate the Review	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.2 Provide Review Feedbacks	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.3 Review and Understand Feedbacks	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.4 Conduct the Review	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.5 Update Test Scripts after Review	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.6 Check Test Scripts for Approval	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.7 Update Test Scripts after Approval Checking	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
3.8 Approve Test Scripts	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4 Roles																										
4.1 SQE	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.3 Team Leader	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.4 Reviewers	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
4.5 Author	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
5 Tools and Techniques																										
5.1 MS Outlook	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
5.2 Custom Action Item DB	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
Process Cluster																										

Figure E.6 Process Similarity Matrix for Test Development Peer Review Process

Metric Usability Questionnaires

- The questionnaires for metrics of test design process are given by Figure E.7 through Figure E.11.
- The questionnaires for metrics of test procedure development process are given by Figure E.12 through Figure E.15.
- The questionnaires for metrics of test development are given by Figure E.16 through Figure E.19.
- The questionnaires for metrics of test development peer review process are given by Figure E.20 through Figure E.25.
- The questionnaires for metrics of overall reviews are given by Figure E.26 and Figure E.27.

Metric Name: Test Design Internal Review Effort		Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Conceptual Definition: Total effort spent for internal reviews of test design of a package		F : Indicatiors of the atribute are fully satisfied (%86-100)		
Assessed On: 02/05/2006		L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed By: A.Tarhan		P : Indicatiors of the atribute are largely satisfied (%16-50)		
		N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes		**Answers**	**Rating**	**Expected Answers**
	Indicators			
Metric Identity		MUF-1	F	
Q1	Which entity does the metric measure?	Human Resource		
Q2	Which attribute of the entity does the metric measure?	Effort		
Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q4	What is the unit of the metric data?	Person-hours		
Q5	What is the type of the metric data? (integer, real, etc.)	Real		
Q6	What is the range of the metric data?	[0.00, 72.00]		
Data Existence		MUF-2	F	
Q7	Is metric data existent?	Yes		Available > 20
Q8	What is the amount of overall observations?	26		
Q9	What is the amount of missing data points?	1		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	After performing internal reviews		
Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
Q14	Who is responsible for recording metric data?	Test Designer, Team Leader		
Q15	Is all metric data recorded by the responsible body?	Yes		Yes
Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability		MUF-4	L	
Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	No		No
Q24	Is metric data recorded precisely?	Yes		Yes
Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)		Yes
Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
Q27	Is metric data analyzed and reported?	No		Yes
Q28	Is metric data analysis results communicated to process performers?	No		Yes
Q29	Is metric data analysis results communicated to management?	No		Yes
Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability				
Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability				
Q32	Is metric data integrable at project level?	Yes		
Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.7 Metric Usability Questionnaire for “Test Design Internal Review Effort” Base Metric of Test Design Process

Please rate each attribute in four scales, based on asnwers to questions as indicators:

F : Indicatiors of the atribute are fully satisfied (%86-100)

L : Indicatiors of the atribute are largely satisfied (%51-85)

P : Indicatiors of the atribute are largely satisfied (%16-50)

N : Indicatiors of the atribute are not satisfied (%0-15)

Metric Name: Number of test cases

Conceptual Definition: Number of test cases in a package

Assessed On: 02/05/2006

Assessed By: A.Tarhan

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	Which entity does the metric measure?	Test Cases		
	Q2	Which attribute of the entity does the metric measure?	Count		
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
	Q4	What is the unit of the metric data?	N/A		
	Q5	What is the type of the metric data? (integer, real, etc.)	Integer		
	Q6	What is the range of the metric data?	[2, 72]		
Data Existence				F	
	Q7	Is metric data existent?	Yes		Available > 20
	Q8	What is the amount of overall observations?	26		
	Q9	What is the amount of missing data points?	1		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	In the middle (after test cases are identified)		
	Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
	Q14	Who is responsible for recording metric data?	Test Designer		
	Q15	Is all metric data recorded by the responsible body?	Yes		Yes
	Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
	Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
	Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
	Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				L	
	Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No		No
	Q24	Is metric data recorded precisely?	Yes		Yes
	Q25	Is metric data collected for a specific purpose?	Yes (for historical data)		Yes
	Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
	Q27	Is metric data analyzed and reported?	No		Yes
	Q28	Is metric data analysis results communicated to process performers?	No		Yes
	Q29	Is metric data analysis results communicated to management?	No		Yes
	Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability					
	Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability					
	Q32	Is metric data integrable at project level?	Yes		
	Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	Rating	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.8 Metric Usability Questionnaire for "Number of Test Cases" Base Metric of Test Design Process

Metric Name: Actual Test Design Effort
Conceptual Definition: Total actual effort spent for test design of a package
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Please rate each attribute in four scales, based on asnwers to questions as indicators:
F : Indicatiors of the atribute are fully satisfied (%86-100)
L : Indicatiors of the atribute are largely satisfied (%51-85)
P : Indicatiors of the atribute are largely satisfied (%16-50)
N : Indicatiors of the atribute are not satisfied (%0-15)

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	Which entity does the metric measure?	Human Resource		
	Q2	Which attribute of the entity does the metric measure?	Effort		
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q4	What is the unit of the metric data?	Person-hours		
	Q5	What is the type of the metric data? (integer, real, etc.)	Real		
	Q6	What is the range of the metric data?	[6.00, 223.00]		
Data Existence				F	
	Q7	Is metric data existent?	Yes		Available > 20
	Q8	What is the amount of overall observations?	26		
	Q9	What is the amount of missing data points?	1		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end		
	Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
	Q14	Who is responsible for recording metric data?	Test Designer		
	Q15	Is all metric data recorded by the responsible body?	Yes		Yes
	Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
	Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
	Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
	Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				L	
	Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No		No
	Q24	Is metric data recorded precisely?	Yes		Yes
	Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)		Yes
	Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
	Q27	Is metric data analyzed and reported?	No		Yes
	Q28	Is metric data analysis results communicated to process performers?	No		Yes
	Q29	Is metric data analysis results communicated to management?	No		Yes
	Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability					
	Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability					
	Q32	Is metric data integrable at project level?	Yes		
	Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.9 Metric Usability Questionnaire for "Actual Test Design Effort" Base Metric of Test Design Process

Metric Name: Test Design Productivity	Please rate each attribute in four scales, based on asnwers to questions as indicators:
Conceptual Definition: Number of test cases designed per hour	F : Indications of the atribute are fully satisfied (%86-100)
Assessed On: 02/05/2006	L : Indications of the atribute are largely satisfied (%51-85)
Assessed By: A.Tarhan	P : Indications of the atribute are largely satisfied (%16-50)
	N : Indications of the atribute are not satisfied (%0-15)

Attributes	Indicators		Answers		Rating	Expected Answers
Metric Identity				MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Number of test cases / Actual test design effort			
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio			Ratio, Absolute
	Q3	What is the unit of the metric data?	#TC / hour			
	Q4	What is the type of the metric data? (integer, real, etc.)	Real			
	Q5	What is the range of the metric data?	[0.08, 2.19]			
Data Existence					F	
	Q6	Is metric data existent?	Yes			Available > 10
	Q7	What is the amount of overall observations?	26			
	Q8	What is the amount of missing data points?	1			
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No			
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.			
Data Verifiability				MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula			
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes			Yes
	Q13	Is all metric data calculated according to metric formula?	Yes			Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet			
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes			Yes
Data Dependability					N/A	
	Q16	Is metric data stored precisely?	N/A			Yes
	Q17	Is metric data stored for a specific purpose?	N/A			Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A			Yes
	Q19	Is metric data analyzed and reported?	N/A			Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A			Yes
	Q21	Is metric data analysis results communicated to management?	N/A			Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A			Yes
Data Normalizability						
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No			
Data Integrability						
	Q24	Is metric data integrable at project level?	Yes			
	Q25	Is metric data integrable at organization level?	No			

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.10 Metric Usability Questionnaire for “Test Design Productivity” Derived Metric of Test Design Process

Metric Name: Percent of Test Design Internal Review Effort		Please rate each attribute in four scales, based on asnwers to questions as indicators:	
Conceptual Definition: Percent of Test Desing Internal Review Effort within Actual Test Design Effort		F : Indications of the atribute are fully satisfied (%86-100)	
Assessed On: 02/05/2006		L : Indications of the atribute are largely satisfied (%51-85)	
Assessed By: A.Tarhan		P : Indications of the atribute are largely satisfied (%16-50)	
		N : Indications of the atribute are not satisfied (%0-15)	

Attributes / Indicators		Answers	Rating	Expected Answers
Metric Identity		MUF-1	F	
Q1	What is the the metric formula? (please refer to related base metrics)	Test design internal review effort / Actual test design effort		
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q3	What is the unit of the metric data?	N/A		
Q4	What is the type of the metric data? (integer, real, etc.)	Real		
Q5	What is the range of the metric data?	[0.00, 0.53]		
Data Existence			F	
Q6	Is metric data existent?	Yes		Available > 10
Q7	What is the amount of overall observations?	26		
Q8	What is the amount of missing data points?	1		
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
Q13	Is all metric data calculated according to metric formula?	Yes		Yes
Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			N/A	
Q16	Is metric data stored precisely?	N/A		Yes
Q17	Is metric data stored for a specific purpose?	N/A		Yes
Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
Q19	Is metric data analyzed and reported?	N/A		Yes
Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
Q21	Is metric data analysis results communicated to management?	N/A		Yes
Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability				
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability				
Q24	Is metric data integrable at project level?	Yes		
Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.11 Metric Usability Questionnaire for “Percent of Test Design Internal Review Effort” Derived Metric of Test Design Process

Metric Name: Actual Test Procedure Development Effort	Please rate each attribute in four scales, based on asnwers to questions as indicators:
Conceptual Definition: Total actual effort spent for test procedure development of a package	F : Indicatiors of the atribute are fully satisfied (%86-100)
Assessed On: 02/05/2006	L : Indicatiors of the atribute are largely satisfied (%51-85)
Assessed By: A.Tarhan	P : Indicatiors of the atribute are largely satisfied (%16-50)
	N : Indicatiors of the atribute are not satisfied (%0-15)

Attributes	Indicators	Answers	Rating	Expected Answers
Metric Identity		MUF-1	F	
Q1	Which entity does the metric measure?	Human Resource		
Q2	Which attribute of the entity does the metric measure?	Effort		
Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q4	What is the unit of the metric data?	Person-hours		
Q5	What is the type of the metric data? (integer, real, etc.)	Real		
Q6	What is the range of the metric data?	[8.30, 232.75]		
Data Existence			F	
Q7	Is metric data existent?	Yes		Available > 20
Q8	What is the amount of overall observations?	52		
Q9	What is the amount of missing data points?	4		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end		
Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
Q14	Who is responsible for recording metric data?	Test Developer		
Q15	Is all metric data recorded by the responsible body?	Yes		Yes
Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			L	
Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	No		No
Q24	Is metric data recorded precisely?	Yes		Yes
Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)		Yes
Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
Q27	Is metric data analyzed and reported?	No		Yes
Q28	Is metric data analysis results communicated to process performers?	No		Yes
Q29	Is metric data analysis results communicated to management?	No		Yes
Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability				
Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability				
Q32	Is metric data integrable at project level?	Yes		
Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	Rating	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.12 Metric Usability Questionnaire for “Actual Test Procedure Development Effort” Base Metric of Test Procedure Development Process

Please rate each attribute in four scales, based on asnwers to questions as indicators:

F : Indicatiors of the atribute are fully satisfied (%86-100)
L : Indicatiors of the atribute are largely satisfied (%51-85)
P : Indicatiors of the atribute are largely satisfied (%16-50)
N : Indicatiors of the atribute are not satisfied (%0-15)

Metric Name: Test Procedure Development Internal Review Effort
Conceptual Definition: Total effort spent for internal reviews of test procedure development of a pack
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Attributes		Answers	Rating	Expected Answers
Indicators				
Metric Identity		MUF-1	F	
Q1	Which entity does the metric measure?	Human Resource		
Q2	Which attribute of the entity does the metric measure?	Effort		
Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q4	What is the unit of the metric data?	Person-hours		
Q5	What is the type of the metric data? (integer, real, etc.)	Real		
Q6	What is the range of the metric data?	[0.00, 58.00]		
Data Existence			F	
Q7	Is metric data existent?	Yes		Available > 20
Q8	What is the amount of overall observations?	52		
Q9	What is the amount of missing data points?	4		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	After performing internal reviews		
Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
Q14	Who is responsible for recording metric data?	Test Designer, Team Leader		
Q15	Is all metric data recorded by the responsible body?	Yes		Yes
Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			L	
Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	No		No
Q24	Is metric data recorded precisely?	Yes		Yes
Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)		Yes
Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
Q27	Is metric data analyzed and reported?	No		Yes
Q28	Is metric data analysis results communicated to process performers?	No		Yes
Q29	Is metric data analysis results communicated to management?	No		Yes
Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability				
Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability				
Q32	Is metric data integrable at project level?	Yes		
Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.13 Metric Usability Questionnaire for “Test Procedure Development Internal Review Effort” Base Metric of Test Procedure Development Process

Please rate each attribute in four scales, based on asnwers to questions as indicators:

F : Indicatiors of the atribute are fully satisfied (%86-100)
L : Indicatiors of the atribute are largely satisfied (%51-85)
P : Indicatiors of the atribute are largely satisfied (%16-50)
N : Indicatiors of the atribute are not satisfied (%0-15)

Metric Name: Percent of Test Procedure Development Internal Review Effort
Conceptual Definition: Percent of Test Procedure Development Internal Review Effort within Actual Test
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Test procedure development internal review effort / Actual test procedure development effort		
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q3	What is the unit of the metric data?	N/A		
	Q4	What is the type of the metric data? (integer, real, etc.)	Real		
	Q5	What is the range of the metric data?	[0.00, 0.39]		
Data Existence				F	
	Q6	Is metric data existent?	Yes		Available > 10
	Q7	What is the amount of overall observations?	52		
	Q8	What is the amount of missing data points?	4		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number		
Data Verifiability			MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
	Q13	Is all metric data calculated according to metric formula?	Yes		Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				N/A	
	Q16	Is metric data stored precisely?	N/A		Yes
	Q17	Is metric data stored for a specific purpose?	N/A		Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
	Q19	Is metric data analyzed and reported?	N/A		Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
	Q21	Is metric data analysis results communicated to management?	N/A		Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability					
	Q24	Is metric data integrable at project level?	Yes		
	Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.14 Metric Usability Questionnaire for “Percent of Test Procedure Development Internal Review Effort” Derived Metric of Test Procedure Development Process

		Please rate each attribute in four scales, based on asnwers to questions as indicators:
Metric Name: Test Procedure Development Productivity		F : Indications of the atribute are fully satisfied (%86-100)
Conceptual Definition: Number of test cases developed per hour		L : Indications of the atribute are largely satisfied (%51-85)
Assessed On: 02/05/2006		P : Indications of the atribute are largely satisfied (%16-50)
Assessed By: A.Tarhan		N : Indications of the atribute are not satisfied (%0-15)

Attributes		Answers		Rating	Expected Answers
	Indicators				
Metric Identity			MUF-1	F	
Q1	What is the the metric formula? (please refer to related base metrics)	Number of test cases / Actual test procedure development effort			
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio			Ratio, Absolute
Q3	What is the unit of the metric data?	#TC/hour			
Q4	What is the type of the metric data? (integer, real, etc.)	Real			
Q5	What is the range of the metric data?	[0.02, 0.84]			
Data Existence				F	
Q6	Is metric data existent?	Yes			Available > 10
Q7	What is the amount of overall observations?	52			
Q8	What is the amount of missing data points?	4			
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No			
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number			
Data Verifiability			MUF-3	F	
Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula			
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes			Yes
Q13	Is all metric data calculated according to metric formula?	Yes			Yes
Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet			
Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes			Yes
Data Dependability				N/A	
Q16	Is metric data stored precisely?	N/A			Yes
Q17	Is metric data stored for a specific purpose?	N/A			Yes
Q18	Is the purpose of metric data storage known by process performers?	N/A			Yes
Q19	Is metric data analyzed and reported?	N/A			Yes
Q20	Is metric data analysis results communicated to process performers?	N/A			Yes
Q21	Is metric data analysis results communicated to management?	N/A			Yes
Q22	Is metric data analysis results used as a basis for decision making?	N/A			Yes
Data Normalizability					
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No			
Data Integrability					
Q24	Is metric data integrable at project level?	Yes			
Q25	Is metric data integrable at organization level?	No			

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.15 Metric Usability Questionnaire for “Test Procedure Development Productivity” Derived Metric of Test Procedure Development Process

		Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Metric Name: Actual Test Development Effort		F : Indications of the atribute are fully satisfied (%86-100)		
Conceptual Definition: Total actual effort spent for test design and test procedure development		L : Indications of the atribute are largely satisfied (%51-85)		
Assessed On: 02/05/2006		P : Indications of the atribute are largely satisfied (%16-50)		
Assessed By: A.Tarhan		N : Indications of the atribute are not satisfied (%0-15)		
Attributes		Answers	Rating	Expected Answers
	Indicators			
Metric Identity		MUF-1	F	
Q1	What is the the metric formula? (please refer to related base metrics)	Actual test design effort + Actual test procedure development effort		
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q3	What is the unit of the metric data?	Person-hours		
Q4	What is the type of the metric data? (integer, real, etc.)	Real		
Q5	What is the range of the metric data?	[11.50, 421.45]		
Data Existence			F	
Q6	Is metric data existent?	Yes		Available > 10
Q7	What is the amount of overall observations?	52		
Q8	What is the amount of missing data points?	4		
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number		
Data Verifiability		MUF-3	F	
Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
Q13	Is all metric data calculated according to metric formula?	Yes		Yes
Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			N/A	
Q16	Is metric data stored precisely?	N/A		Yes
Q17	Is metric data stored for a specific purpose?	N/A		Yes
Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
Q19	Is metric data analyzed and reported?	N/A		Yes
Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
Q21	Is metric data analysis results communicated to management?	N/A		Yes
Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability				
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability				
Q24	Is metric data integrable at project level?	Yes		
Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.16 Metric Usability Questionnaire for “Actual Test Development Effort” Derived Metric of Test Development

Please rate each attribute in four scales, based on asnwers to questions as indicators:

F : Indicatiors of the atribute are fully satisfied (%86-100)
L : Indicatiors of the atribute are largely satisfied (%51-85)
P : Indicatiors of the atribute are largely satisfied (%16-50)
N : Indicatiors of the atribute are not satisfied (%0-15)

Metric Name: Percent of Test Development Internal Review Effort
Conceptual Definition: Percent of Test Development Internal Review Effort within Actual Test Developme
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Test development internal review effort / Actual test development effort		
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q3	What is the unit of the metric data?	N/A		
	Q4	What is the type of the metric data? (integer, real, etc.)	Real		
	Q5	What is the range of the metric data?	[0.00, 0.39]		
Data Existence				F	
	Q6	Is metric data existent?	Yes		Available > 10
	Q7	What is the amount of overall observations?	52		
	Q8	What is the amount of missing data points?	4		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number		
Data Verifiability			MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
	Q13	Is all metric data calculated according to metric formula?	Yes		Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				N/A	
	Q16	Is metric data stored precisely?	N/A		Yes
	Q17	Is metric data stored for a specific purpose?	N/A		Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
	Q19	Is metric data analyzed and reported?	N/A		Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
	Q21	Is metric data analysis results communicated to management?	N/A		Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability					
	Q24	Is metric data integrable at project level?	Yes		
	Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.17 Metric Usability Questionnaire for "Percent of Test Development Internal Review Effort" Derived Metric of Test Development

Metric Name: Test Development Internal Review Effort
Conceptual Definition: Total actual effort spent for internal reviews of test design and development
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Please rate each attribute in four scales, based on asnwers to questions as indicators:
F : Indications of the atribute are fully satisfied (%86-100)
L : Indications of the atribute are largely satisfied (%51-85)
P : Indications of the atribute are largely satisfied (%16-50)
N : Indications of the atribute are not satisfied (%0-15)

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Test design internal review effort + Test procedure development internal review effort		
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q3	What is the unit of the metric data?	Person-hours		
	Q4	What is the type of the metric data? (integer, real, etc.)	Real		
	Q5	What is the range of the metric data?	[0.0, 89.50]		
Data Existence				F	
	Q6	Is metric data existent?	Yes		Available > 10
	Q7	What is the amount of overall observations?	52		
	Q8	What is the amount of missing data points?	4		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number		
Data Verifiability			MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
	Q13	Is all metric data calculated according to metric formula?	Yes		Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				N/A	
	Q16	Is metric data stored precisely?	N/A		Yes
	Q17	Is metric data stored for a specific purpose?	N/A		Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
	Q19	Is metric data analyzed and reported?	N/A		Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
	Q21	Is metric data analysis results communicated to management?	N/A		Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability					
	Q24	Is metric data integrable at project level?	Yes		
	Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.18 Metric Usability Questionnaire for "Test Development Internal Review Effort" Derived Metric of Test Development

		Please rate each attribute in four scales, based on asnwers to questions as indicators:
Metric Name: Test Development Productivity		F : Indications of the atribute are fully satisfied (%86-100)
Conceptual Definition: Number of test cases completed per hour		L : Indications of the atribute are largely satisfied (%51-85)
Assessed On: 02/05/2006		P : Indications of the atribute are largely satisfied (%16-50)
Assessed By: A.Tarhan		N : Indications of the atribute are not satisfied (%0-15)

Attributes		Answers	Rating	Expected Answers
Indicators				
Metric Identity		MUF-1	F	
Q1	What is the the metric formula? (please refer to related base metrics)	Number of test cases / Actual test development effort		
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q3	What is the unit of the metric data?	#TC/hour		
Q4	What is the type of the metric data? (integer, real, etc.)	Real		
Q5	What is the range of the metric data?	[0.02, 0.31]		
Data Existence			F	
Q6	Is metric data existent?	Yes		Available > 10
Q7	What is the amount of overall observations?	52		
Q8	What is the amount of missing data points?	4		
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number		
Data Verifiability		MUF-3	F	
Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
Q13	Is all metric data calculated according to metric formula?	Yes		Yes
Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			N/A	
Q16	Is metric data stored precisely?	N/A		Yes
Q17	Is metric data stored for a specific purpose?	N/A		Yes
Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
Q19	Is metric data analyzed and reported?	N/A		Yes
Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
Q21	Is metric data analysis results communicated to management?	N/A		Yes
Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability				
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability				
Q24	Is metric data integrable at project level?	Yes		
Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.19 Metric Usability Questionnaire for “Test Development Productivity” Derived Metric of Test Development

Metric Name: Number of action items	Please rate each attribute in four scales, based on asnwers to questions as indicators:
Conceptual Definition: Number of action items detected in a test development peer review	F : Indicatiors of the atribute are fully satisfied (%86-100)
Assessed On: 02/05/2006	L : Indicatiors of the atribute are largely satisfied (%51-85)
Assessed By: A.Tarhan	P : Indicatiors of the atribute are largely satisfied (%16-50)
	N : Indicatiors of the atribute are not satisfied (%0-15)

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	Which entity does the metric measure?	Action Items		
	Q2	Which attribute of the entity does the metric measure?	Count		
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Absolute		Ratio, Absolute
	Q4	What is the unit of the metric data?	N/A		
	Q5	What is the type of the metric data? (integer, real, etc.)	Integer		
	Q6	What is the range of the metric data?	[1, 56]		
Data Existence				F	
	Q7	Is metric data existent?	Yes		Available > 20
	Q8	What is the amount of overall observations?	31		
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the peer review		
	Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
	Q14	Who is responsible for recording metric data?	SQE		
	Q15	Is all metric data recorded by the responsible body?	Yes		Yes
	Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
	Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
	Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
	Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				L	
	Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No		No
	Q24	Is metric data recorded precisely?	Yes		Yes
	Q25	Is metric data collected for a specific purpose?	Yes (for historical data)		Yes
	Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
	Q27	Is metric data analyzed and reported?	No		Yes
	Q28	Is metric data analysis results communicated to process performers?	No		Yes
	Q29	Is metric data analysis results communicated to management?	No		Yes
	Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability					
	Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability					
	Q32	Is metric data integrable at project level?	Yes		
	Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	Rating	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.20 Metric Usability Questionnaire for “Number of Action Items” Base Metric of Test Development Peer Review Process

		Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Metric Name: Test Development Peer Review Effort		F : Indicatiors of the atribute are fully satisfied (%86-100)		
Conceptual Definition: Total effort spent for review in a test development peer review		L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed On: 02/05/2006		P : Indicatiors of the atribute are largely satisfied (%16-50)		
Assessed By: A.Tarhan		N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes		Answers	Rating	Expected Answers
	Indicators			
Metric Identity		MUF-1	F	
Q1	Which entity does the metric measure?	Human Resource		
Q2	Which attribute of the entity does the metric measure?	Effort		
Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q4	What is the unit of the metric data?	Person-hours		
Q5	What is the type of the metric data? (integer, real, etc.)	Real		
Q6	What is the range of the metric data?	[2.60, 35.00]		
Data Existence			F	
Q7	Is metric data existent?	Yes		Available > 20
Q8	What is the amount of overall observations?	31		
Q9	What is the amount of missing data points?	0		
Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of the peer review		
Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
Q14	Who is responsible for recording metric data?	SQE, Reviewers		
Q15	Is all metric data recorded by the responsible body?	Yes		Yes
Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			L	
Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
Q23	Are the frequencies for data generation, recording, and storing different?	No		No
Q24	Is metric data recorded precisely?	Yes		Yes
Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)		Yes
Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
Q27	Is metric data analyzed and reported?	No		Yes
Q28	Is metric data analysis results communicated to process performers?	No		Yes
Q29	Is metric data analysis results communicated to management?	No		Yes
Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability				
Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of action items detected in peer review)		
Data Integrability				
Q32	Is metric data integrable at project level?	Yes		
Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.21 Metric Usability Questionnaire for “Test Development Peer Review Effort” Base Metric of Test Development Peer Review Process

Metric Name: Test Development Peer Review Update Effort			Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Conceptual Definition: Effort spent for updating action items detected in a test development peer revi			F : Indicatiors of the atribute are fully satisfied (%86-100)		
Assessed On: 02/05/2006			L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed By: A.Tarhan			P : Indicatiors of the atribute are largely satisfied (%16-50)		
			N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes			Answers	Rating	Expected Answers
	Indicators				
Metric Identity			MUF-1	F	
	Q1	Which entity does the metric measure?	Human Resource		
	Q2	Which attribute of the entity does the metric measure?	Effort		
	Q3	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q4	What is the unit of the metric data?	Person-hours		
	Q5	What is the type of the metric data? (integer, real, etc.)	Real		
	Q6	What is the range of the metric data?	[0.00, 27.00]		
Data Existence				F	
	Q7	Is metric data existent?	Yes		Available > 20
	Q8	What is the amount of overall observations?	31		
	Q9	What is the amount of missing data points?	0		
	Q10	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q11	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q12	When is metric data recorded in the process? (at start, middle, end, later, etc.)	At the end of update (before closing the review)		
	Q13	Is all metric data recorded at the same place in the process? (at start, middle, end, later,	Yes		Yes
	Q14	Who is responsible for recording metric data?	Author		
	Q15	Is all metric data recorded by the responsible body?	Yes		Yes
	Q16	How is metric data recorded? (on a form, report, tool, etc.)	On an excel sheet		
	Q17	Is all metric data recorded the same way? (on a form, report, tool, etc.)	Yes		Yes
	Q18	Where is metric data stored? (in a file, database, etc.)	In a file		
	Q19	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				L	
	Q20	What is the frequency of generating metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q21	What is the frequency of recording metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q22	What is the frequency of storing metric data? (asynchronously, daily, weekly, monthly,	Asynchronously		
	Q23	Are the frequencies for data generation, recording, and storing different?	No		No
	Q24	Is metric data recorded precisely?	Yes		Yes
	Q25	Is metric data collected for a specific purpose?	Yes (for effort tracking)		Yes
	Q26	Is the purpose of metric data collection known by process performers?	Yes		Yes
	Q27	Is metric data analyzed and reported?	No		Yes
	Q28	Is metric data analysis results communicated to process performers?	No		Yes
	Q29	Is metric data analysis results communicated to management?	No		Yes
	Q30	Is metric data analysis results used as a basis for decision making?	No		Yes
Data Normalizability					
	Q31	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of action items detected in a peer review)		
Data Integrability					
	Q32	Is metric data integrable at project level?	Yes		
	Q33	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUA-1)	F	F
Data Existence (MUA-2)	F	F
Data Verifiability (MUA-3)	F	L or F
Data Dependability (MUA-4)	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.22 Metric Usability Questionnaire for “Test Development Peer Review Update Effort” Base Metric of Test Development Peer Review Process

Metric Name: Action Item Density	Please rate each attribute in four scales, based on asnwers to questions as indicators:
Conceptual Definition: Number of action items detected per test case in a peer review	F : Indications of the atribute are fully satisfied (%86-100)
Assessed On: 02/05/2006	L : Indications of the atribute are largely satisfied (%51-85)
Assessed By: A.Tarhan	P : Indications of the atribute are largely satisfied (%16-50)
	N : Indications of the atribute are not satisfied (%0-15)

Attributes		Answers	Rating	Expected Answers
Indicators				
Metric Identity		MUF-1	F	
Q1	What is the the metric formula? (please refer to related base metrics)	Number of action items / Number of test cases		
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q3	What is the unit of the metric data?	#AI / #TC		
Q4	What is the type of the metric data? (integer, real, etc.)	Real		
Q5	What is the range of the metric data?	[0.17, 30.00]		
Data Existence			F	
Q6	Is metric data existent?	Yes		Available > 10
Q7	What is the amount of overall observations?	31		
Q8	What is the amount of missing data points?	0		
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
Q13	Is all metric data calculated according to metric formula?	Yes		Yes
Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			N/A	
Q16	Is metric data stored precisely?	N/A		Yes
Q17	Is metric data stored for a specific purpose?	N/A		Yes
Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
Q19	Is metric data analyzed and reported?	N/A		Yes
Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
Q21	Is metric data analysis results communicated to management?	N/A		Yes
Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability				
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability				
Q24	Is metric data integrable at project level?	Yes		
Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.23 Metric Usability Questionnaire for “Action Item Density” Derived Metric of Test Development Peer Review Process

			Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Metric Name: Action Item Detection Efficiency			F : Indicatiors of the atribute are fully satisfied (%86-100)		
Conceptual Definition: Number ofaction items detected per hour of peer review			L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed On: 02/05/2006			P : Indicatiors of the atribute are largely satisfied (%16-50)		
Assessed By: A.Tarhan			N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes			Answers	Rating	Expected Answers
	Indicators				
Metric Identity			MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Number of action items / Test development peer review effort		
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q3	What is the unit of the metric data?	#TC / hour		
	Q4	What is the type of the metric data? (integer, real, etc.)	Real		
	Q5	What is the range of the metric data?	[0.21, 2.98]		
Data Existence				F	
	Q6	Is metric data existent?	Yes		Available > 10
	Q7	What is the amount of overall observations?	31		
	Q8	What is the amount of missing data points?	0		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
	Q13	Is all metric data calculated according to metric formula?	Yes		Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				N/A	
	Q16	Is metric data stored precisely?	N/A		Yes
	Q17	Is metric data stored for a specific purpose?	N/A		Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
	Q19	Is metric data analyzed and reported?	N/A		Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
	Q21	Is metric data analysis results communicated to management?	N/A		Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability					
	Q24	Is metric data integrable at project level?	Yes		
	Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.24 Metric Usability Questionnaire for “Action Item Detection Efficiency” Derived Metric of Test Development Peer Review Process

Metric Name: Action Item Resolution Efficiency		Please rate each attribute in four scales, based on asnwers to questions as indicators:
Conceptual Definition: Number of action items resolved per hour of update effort spent after a peer review		F : Indications of the atribute are fully satisfied (%86-100)
Assessed On: 02/05/2006		L : Indications of the atribute are largely satisfied (%51-85)
Assessed By: A.Tarhan		P : Indications of the atribute are largely satisfied (%16-50)
		N : Indications of the atribute are not satisfied (%0-15)

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Number of action items / Test development peer review update effort		
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q3	What is the unit of the metric data?	#AI / hour		
	Q4	What is the type of the metric data? (integer, real, etc.)	Real		
	Q5	What is the range of the metric data?	[0.31, 112.00]		
Data Existence				F	
	Q6	Is metric data existent?	Yes		Available > 10
	Q7	What is the amount of overall observations?	31		
	Q8	What is the amount of missing data points?	1		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
	Q13	Is all metric data calculated according to metric formula?	Yes		Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				N/A	
	Q16	Is metric data stored precisely?	N/A		Yes
	Q17	Is metric data stored for a specific purpose?	N/A		Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
	Q19	Is metric data analyzed and reported?	N/A		Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
	Q21	Is metric data analysis results communicated to management?	N/A		Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability					
	Q24	Is metric data integrable at project level?	Yes		
	Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.25 Metric Usability Questionnaire for “Action Item Resolution Efficiency” Derived Metric of Test Development Peer Review Process

Please rate each attribute in four scales, based on asnwers to questions as indicators:
F : Indications of the atribute are fully satisfied (%86-100)
L : Indications of the atribute are largely satisfied (%51-85)
P : Indications of the atribute are largely satisfied (%16-50)
N : Indications of the atribute are not satisfied (%0-15)

Metric Name: Total Review Effort
Conceptual Definition: Effort spent for peer review and internal reviews of test development
Assessed On: 02/05/2006
Assessed By: A.Tarhan

Attributes	Indicators		Answers	Rating	Expected Answers
Metric Identity			MUF-1	F	
	Q1	What is the the metric formula? (please refer to related base metrics)	Test Development Internal Review Effort + Test Development Review Effort		
	Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
	Q3	What is the unit of the metric data?	person-hour		
	Q4	What is the type of the metric data? (integer, real, etc.)	Real		
	Q5	What is the range of the metric data?	[9.50, 123.50]		
Data Existence				F	
	Q6	Is metric data existent?	Yes		Available > 10
	Q7	What is the amount of overall observations?	31		
	Q8	What is the amount of missing data points?	0		
	Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
	Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability			MUF-3	F	
	Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
	Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
	Q13	Is all metric data calculated according to metric formula?	Yes		Yes
	Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
	Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability				N/A	
	Q16	Is metric data stored precisely?	N/A		Yes
	Q17	Is metric data stored for a specific purpose?	N/A		Yes
	Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
	Q19	Is metric data analyzed and reported?	N/A		Yes
	Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
	Q21	Is metric data analysis results communicated to management?	N/A		Yes
	Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability					
	Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	Yes (number of test cases in the package)		
Data Integrability					
	Q24	Is metric data integrable at project level?	Yes		
	Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.26 Metric Usability Questionnaire for “Total Review Effort” Derived Metric of Overall Reviews

Metric Name: Total Review Effort per Test Case		Please rate each attribute in four scales, based on asnwers to questions as indicators:		
Conceptual Definition: Effort spent for peer and internal reviews of a package per test case		F : Indicatiors of the atribute are fully satisfied (%86-100)		
Assessed On: 02/05/2006		L : Indicatiors of the atribute are largely satisfied (%51-85)		
Assessed By: A.Tarhan		P : Indicatiors of the atribute are largely satisfied (%16-50)		
		N : Indicatiors of the atribute are not satisfied (%0-15)		
Attributes		Answers	Rating	Expected Answers
	Indicators			
Metric Identity		MUF-1	F	
Q1	What is the the metric formula? (please refer to related base metrics)	Total review effort / Number of test cases		
Q2	What is the scale of the metric data? (nominal, ordinal, interval, ratio, absolute)	Ratio		Ratio, Absolute
Q3	What is the unit of the metric data?	hour / #TC		
Q4	What is the type of the metric data? (integer, real, etc.)	Real		
Q5	What is the range of the metric data?	[0.43, 30.50]		
Data Existence			F	
Q6	Is metric data existent?	Yes		Available > 10
Q7	What is the amount of overall observations?	31		
Q8	What is the amount of missing data points?	0		
Q9	Are data points missing in periods? (If yes, please state observation numbers for missing	No		
Q10	Is metric data time sequenced? (If no, please state how metric data is sequenced)	No. Sequenced by package number.		
Data Verifiability		MUF-3	F	
Q11	How is metric data calculated? (by a tool, manually, etc.)	On excel sheet, by automatic formula		
Q12	Is all metric data calculated the same way? (by a tool, manually, etc.)	Yes		Yes
Q13	Is all metric data calculated according to metric formula?	Yes		Yes
Q14	Where is metric data stored? (in a file, database, etc.)	In excel sheet		
Q15	Is all metric data stored in the same place? (in a file, database, etc.)	Yes		Yes
Data Dependability			N/A	
Q16	Is metric data stored precisely?	N/A		Yes
Q17	Is metric data stored for a specific purpose?	N/A		Yes
Q18	Is the purpose of metric data storage known by process performers?	N/A		Yes
Q19	Is metric data analyzed and reported?	N/A		Yes
Q20	Is metric data analysis results communicated to process performers?	N/A		Yes
Q21	Is metric data analysis results communicated to management?	N/A		Yes
Q22	Is metric data analysis results used as a basis for decision making?	N/A		Yes
Data Normalizability				
Q23	Can metric data be normalized by parameters or metrics? (If yes, please specify them)	No		
Data Integrability				
Q24	Is metric data integrable at project level?	Yes		
Q25	Is metric data integrable at organization level?	No		

Metric Usability Attributes	**Rating**	Expected Rating
Metric Identity (MUF-1)	F	F
Data Existence (MUF-2)	F	F
Data Verifiability (MUF-3)	F	L or F
Data Dependability (MUF-4)	N/A	L or F
MUF-3&4 for base metric-1	L	L or F
MUF-3&4 for base metric-2	L	L or F
Metric Usability Result	**L**	L or F (Usable) -- Not Usable otherwise

Figure E.27 Metric Usability Questionnaire for “Total Review Effort per Test Case” Derived Metric of Overall Reviews

Process Execution Questionnaires

- The questionnaire for the only out-of control point of test design process is given by Figure E.28.
- The questionnaires for out-of control points of test procedure development process are given by Figure E.29 through Figure E.31.
- The questionnaires for out-of control points of test development process are given by Figure E.32 and Figure E.33.
- The questionnaire for the only out-of control point of test development peer review process is given by Figure E.34.
- The questionnaires for out-of control points of overall reviews are given by Figure E.35 and Figure E.36.

Process Name: Test Design		Recorded On: 21.05.2006	
Process Execution No: 11		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	The package was designed by a very experienced test designer who already has domain knowledge.
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.28 Process Execution Questionnaire for Test Design PE # 11

Process Name: Test Script Development		Recorded On: 21.05.2006	
Process Execution No: 1		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	No	The first package under development
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The package has passed many revisions, it was much like an example package. It enabled on-the-job training.		

Figure E.29 Process Execution Questionnaire for Test Proc. Development PE # 1

Process Name: Test Procedure Development		Recorded On: 21.05.2006	
Process Execution No: 11		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	The package was designed by a very experienced test script developer who already has domain knowledge.
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.30 Process Execution Questionnaire for Test Proc. Development PE # 11

Process Name: Test Script Development		Recorded On: 21.05.2006	
Process Execution No: 45		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.31 Process Execution Questionnaire for Test Proc. Development PE # 45

Process Name: Test Development		Recorded On: 21.05.2006	
Process Execution No: 11		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	The package was designed by a very experienced test developer who already has domain knowledge.
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.32 Process Execution Questionnaire for Test Development PE # 11

Process Name: Test Procedure Development		Recorded On: 21.05.2006	
Process Execution No: 45		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.33 Process Execution Questionnaire for Test Development PE # 45

Process Name: Test Development Peer Review		Recorded On: 21.05.2005	
Process Execution No: 44		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			
	The package under review was developed by a very inexperienced staff, and high defectiveness was not a surprise.		

Figure E.34 Process Execution Questionnaire for Test Development Peer Review PE # 44

Process Name: Overall Reviews		Recorded On: 21.05.2006	
Process Execution No: 14		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.35 Process Execution Questionnaire for Overall Reviews PE # 14

Process Name: Overall Reviews		Recorded On: 21.05.2006	
Process Execution No: 24		Recorded By: Ayça Tarhan	
External Attributes		**Status (Yes/No)**	**Explanation**
PROCESS PERFORMERS			
Q1	Are process performers trained in their roles in the process?	Yes	
Q2	Are process performers experienced in their roles in the process?	Yes	
Q3	Are process performers differed per role basis during execution of the process?	No	
PROCESS ENVIRONMENT			
Q4	Has there been a recent change in location?	No	
Q5	Has there been a recent change in support systems? (infrastructure, technology, etc.)	No	
Q6	Has there been a recent change in communication channels and mechanisms? (structure, media, etc.)	No	
Q7	Has there been a recent change in funding and resources allocated for the process?	No	
Q8	Has the process been tailored for this specific execution?	No	
OTHER FACTORS (Please list if any)			

Figure E.36 Process Execution Questionnaire for Overall Reviews PE # 24

Process Attributes Descriptions

- The descriptions for the clusters of test design process are given by Figure E.37 and Figure E.38.
- The description for the only cluster of test procedure development process is given by Figure E.39.
- The description for the only cluster of test development peer review process is given by Figure E.40.

SPC[TM]

Process Attributes Description

Process Name:	Test Design	Described On:	02/06/2006
Process Cluster:	1	Described By:	AT

1. **Inputs**: Please list the inputs to the process.

No	Name	Description
1	Product requirements	

2. **Outputs**: Please list the outputs from the process.

No	Name	Description
1	State Transition Graph	
2	Test Case Document	

3. **Activities (in sequence)**: Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Analyze HL Requirements	
2	Generate Test Design	
3	Identify and Document Test Cases	
4	Team Lead Review	
5	Internal Peer Review	
6	Update Test Design after IP Review	

4. **Roles**: Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Test Designer	
2	Team Leader	
3	External Developer (Peer)	

5. **Tools and Techniques**: Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	MS Visio	
2	MS Excel	
3	Text Editor	

Figure E.37 Process Attributes Description for Test Design Process Cluster 1

SPC[NM]
Process Attributes Description

Process Name:	Test Design	Described On:	02/06/2006
Process Cluster:	2	Described By:	AT

1. **Inputs**: Please list the inputs to the process.

No	Name	Description
1	Product requirements	

2. **Outputs**: Please list the outputs from the process.

No	Name	Description
1	Requirements Tree	
2	Cause-Effect Graph	

3. **Activities (in sequence)**: Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Analyze HL Requirements	
2	Generate Test Design	
3	Identify and Document Test Cases	
4	Team Lead Review	
5	Internal Peer Review	
6	Update Test Design after IP Review	

4. **Roles**: Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Test Designer	
2	Team Leader	
3	External Developer (Peer)	

5. **Tools and Techniques**: Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	MS Visio	
2	Text Editor	

Figure E.38 Process Attributes Description for Test Design Process Cluster 2

SPC

Process Attributes Description

Process Name:	Test Procedure Development	Described On:	02/06/2006
Process Cluster:	Original	Described By:	AT

1. **Inputs**: Please list the inputs to the process.

No	Name	Description
1	Test Case Document	

2. **Outputs**: Please list the outputs from the process.

No	Name	Description
1	Initial Conditions Script	
2	Test Case Script	
3	Traceability Data	

3. **Activities (in sequence)**: Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Prepare Initial Conditions	
2	Develop Test Script	
3	Perform Engineering Run	
4	Update Test Script after Engineering Run	
5	Generate Traceability Data	
6	Team Lead Review	
7	Internal Peer Review	
8	Update Test Script after IP Review	

4. **Roles**: Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	Test Designer	
2	Team Leader	
3	External Developer (Peer)	

5. **Tools and Techniques**: Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	Text Editor	
2	Custom Test Tool	
3	Telelogic DOORS	

Figure E.39 Process Attributes Description for Test Procedure Development Process

SPC
Process Attributes Description

Process Name:	Test Development Peer Review	Described On:	02/06/2006
Process Cluster:	Original	Described By:	AT

1. **Inputs**: Please list the inputs to the process.

No	Name	Description
1	Test Case Scripts	
2	Product Requirements	

2. **Outputs**: Please list the outputs from the process.

No	Name	Description
1	Meeting Request	
2	Action Items	
3	Peer Review Report	
4	Baseline Test Case Scripts	

3. **Activities (in sequence)**: Please list in sequence the activities that are performed while executing the process. You can refer to another process description if an activity consists of sub-activities.

No	Name	Description
1	Initiate the Review	
2	Provide Review Feedbacks	
3	Review and Understand Feedbacks	
4	Conduct the Review	
5	Update Test Scripts after Review	
6	Check Test Scripts for Approval	
7	Update Test Scripts after Approval Checking	
8	Approve Test Scripts	

4. **Roles**: Please list the roles that are allocated responsibilities in the process, by providing references to the activities specified in (3).

No	Name	Description
1	SQE	
2	Team Leader	
3	Reviewers	
4	Author	

5. **Tools and Techniques**: Please list the tools and techniques that are used to support process execution, by providing references to the activities specified in (3).

No	Name	Description
1	MS Outlook	
2	Custom Action Item DB	

Figure E.40 Process Attributes Description for Test Development Peer Review Process

Control Charts

Test Design Process - Productivity

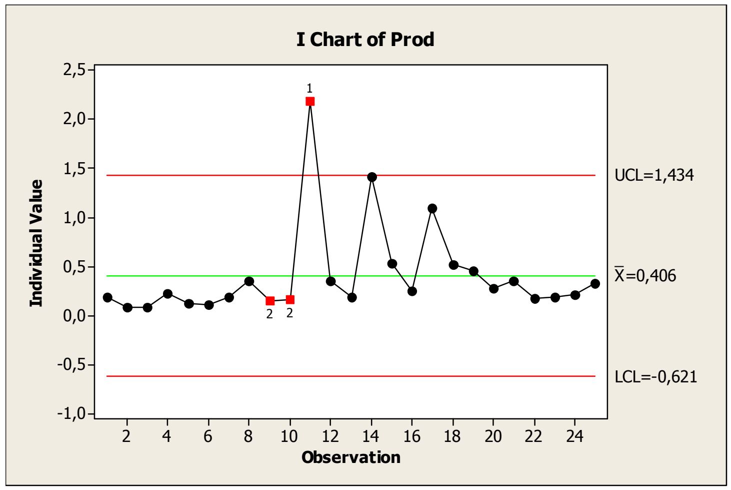


Figure E.41 Control Chart for Productivity of Test Design Process (Overall)

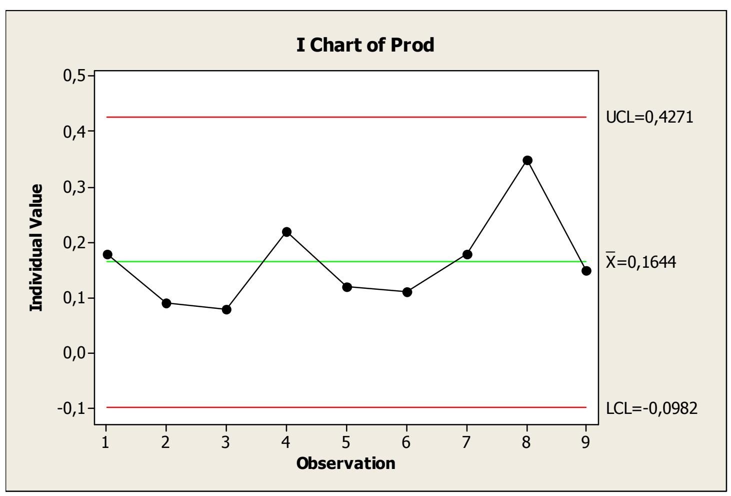


Figure E.42 Control Chart for Productivity of Test Design Process Cluster-1

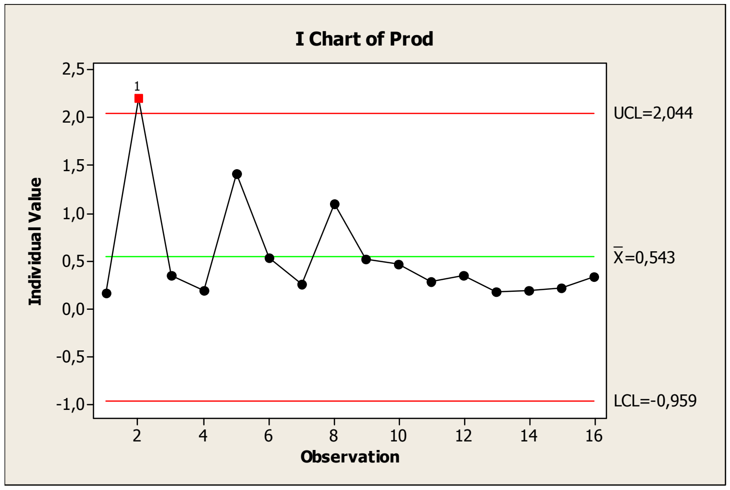


Figure E.43 Control Chart for Productivity of Test Design Process Cluster-2

Test Design Process - Percent of Internal Review Effort

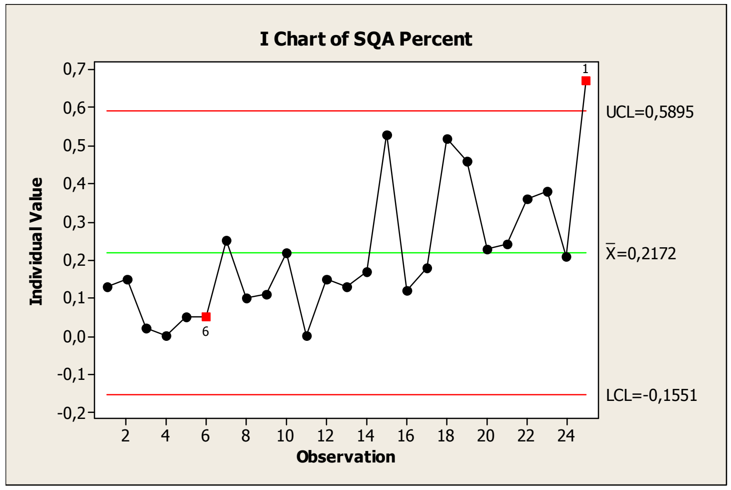


Figure E.44 Control Chart for Percent of Test Design Internal Review Effort of Test Design Process (Overall)

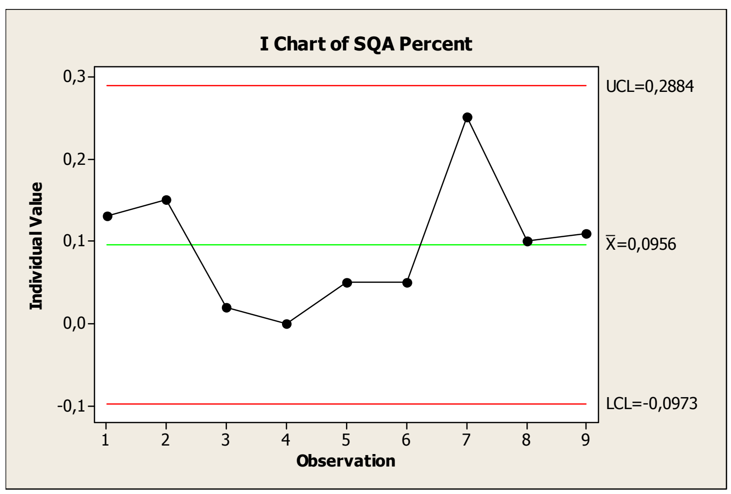


Figure E.45 Control Chart for Percent of Test Design Internal Review Effort of Test Design Process Cluster-1

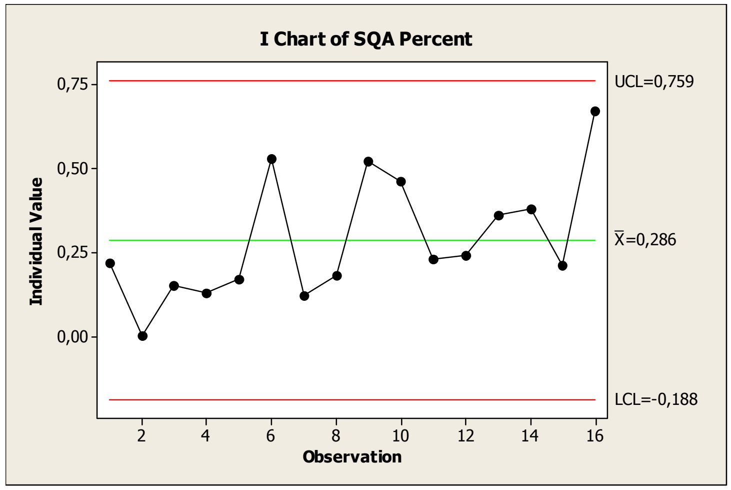


Figure E.46 Control Chart for Percent of Test Design Internal Review Effort of Test Design Process Cluster-2

Test Procedure Development Process - Productivity

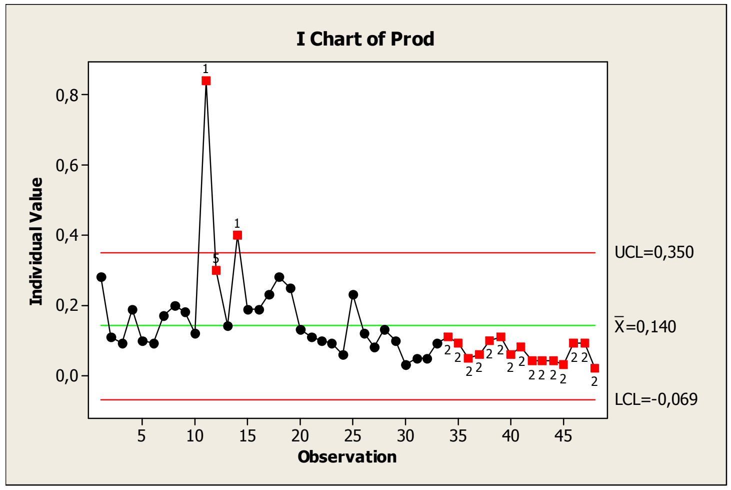


Figure E.47 Control Chart for Productivity of Test Procedure Development Process (Overall)

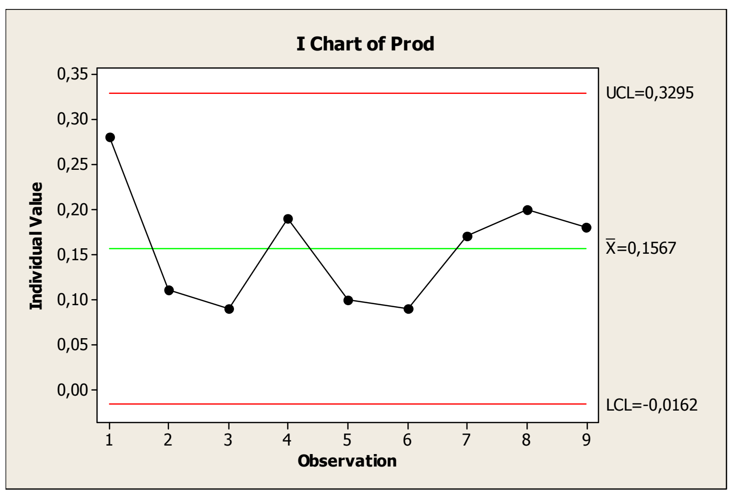


Figure E.48 Control Chart for Productivity of Test Procedure Development Process Cluster-1

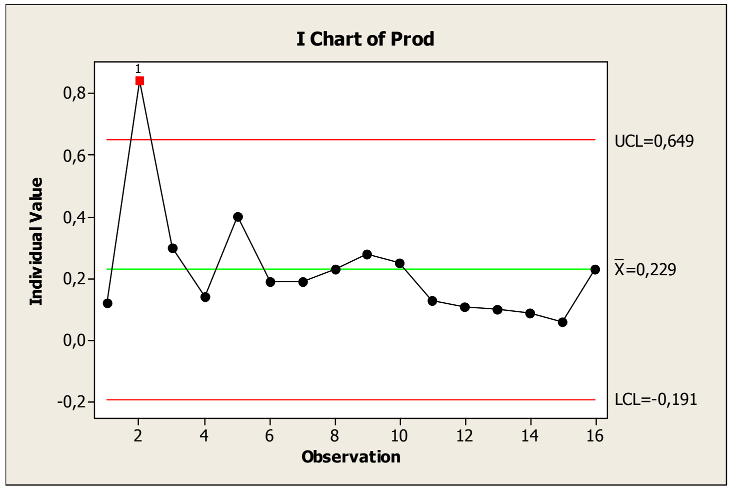


Figure E.49 Control Chart for Productivity of Test Procedure Development Process Cluster-2

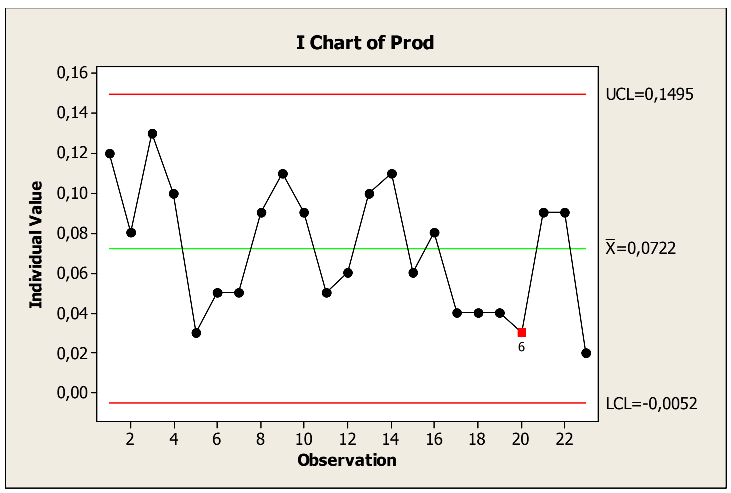


Figure E.50 Control Chart for Productivity of Test Procedure Development Process Cluster-3

Test Procedure Development Process - Percent of Internal Review Effort

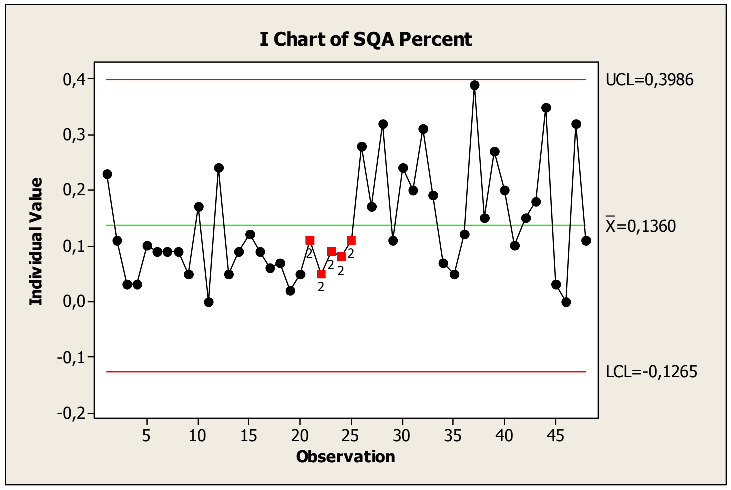


Figure E.51 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process (Overall)

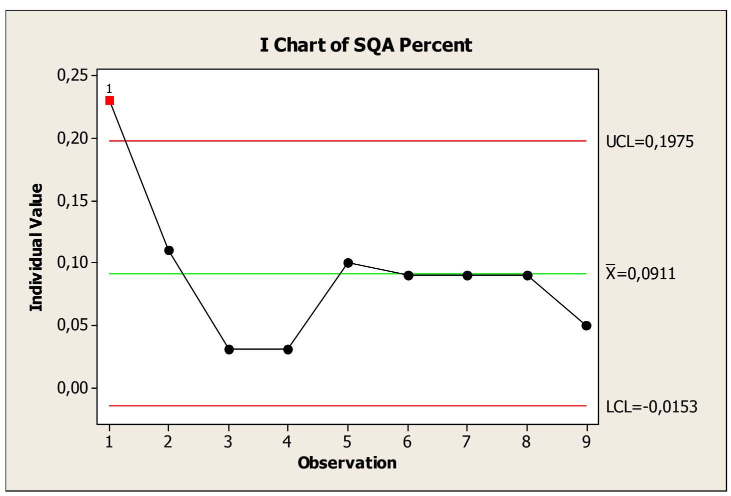


Figure E.52 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process Cluster-1

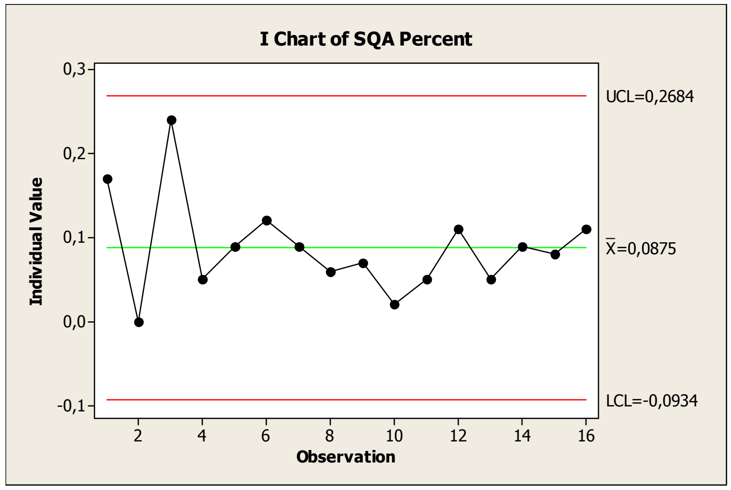


Figure E.53 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process Cluster-2

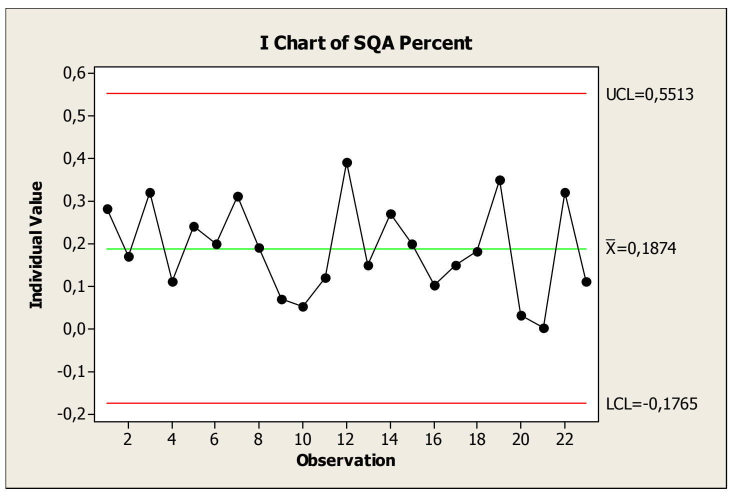


Figure E.54 Control Chart for Percent of Test Design Internal Review Effort of Test Procedure Development Process Cluster-3

Test Development - Productivity

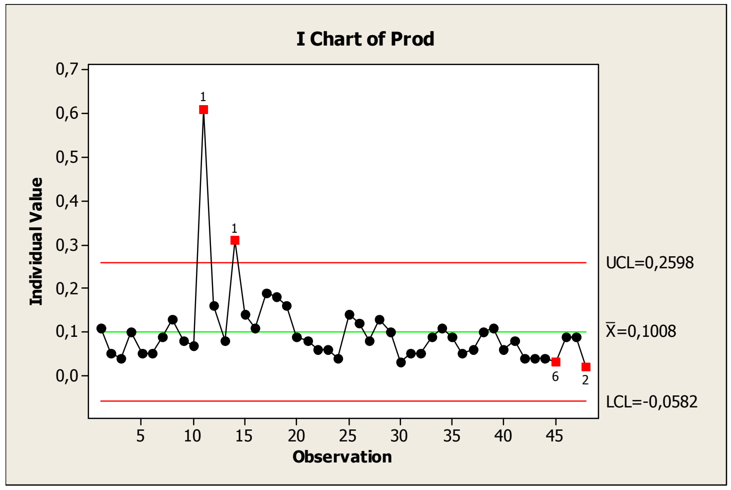


Figure E.55 Control Chart for Productivity of Test Development (Overall)

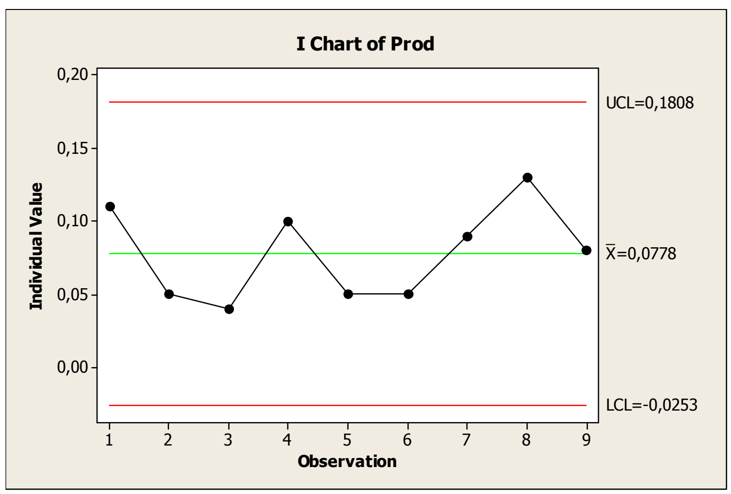


Figure E.56 Control Chart for Productivity of Test Development Cluster-1

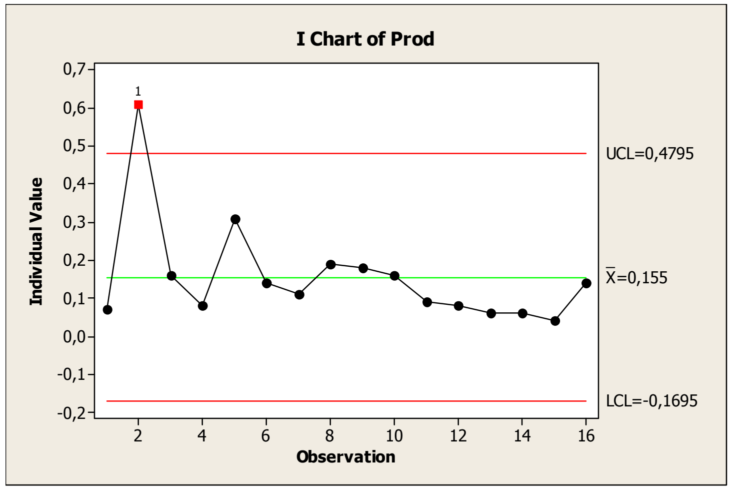


Figure E.57 Control Chart for Productivity of Test Development Cluster-2

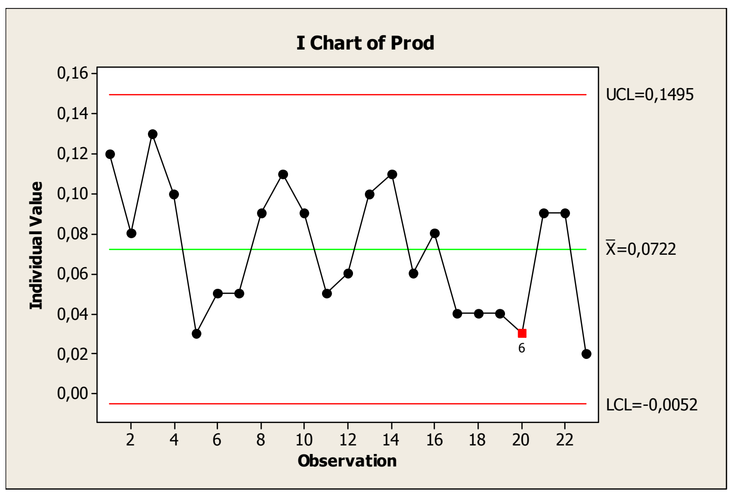


Figure E.58 Control Chart for Productivity of Test Development Cluster-3

Test Procedure Development Process - Percent of Internal Review Effort

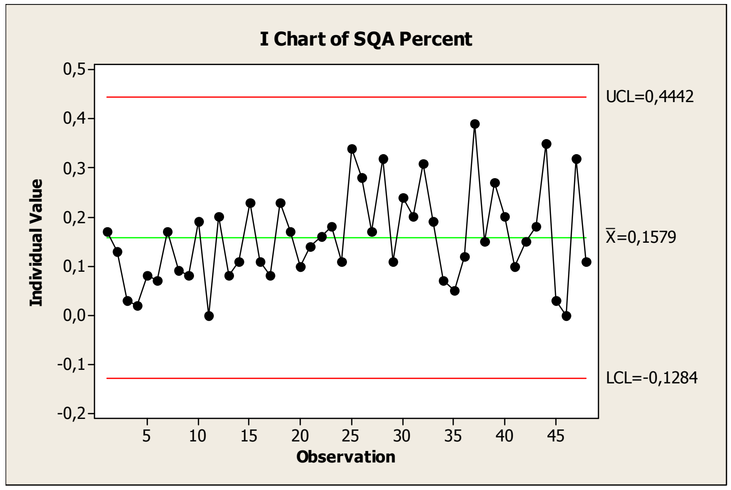


Figure E.59 Control Chart for Percent of Test Design Internal Review Effort of Test Development (Overall)

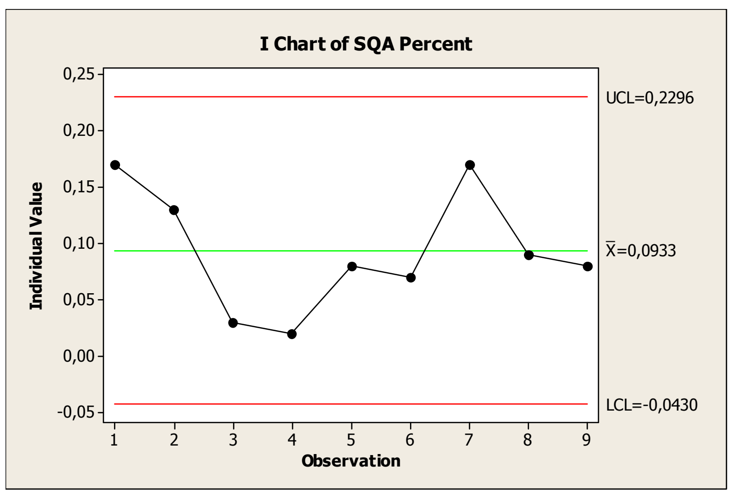


Figure E.60 Control Chart for Percent of Test Design Internal Review Effort of Test Development Cluster-1

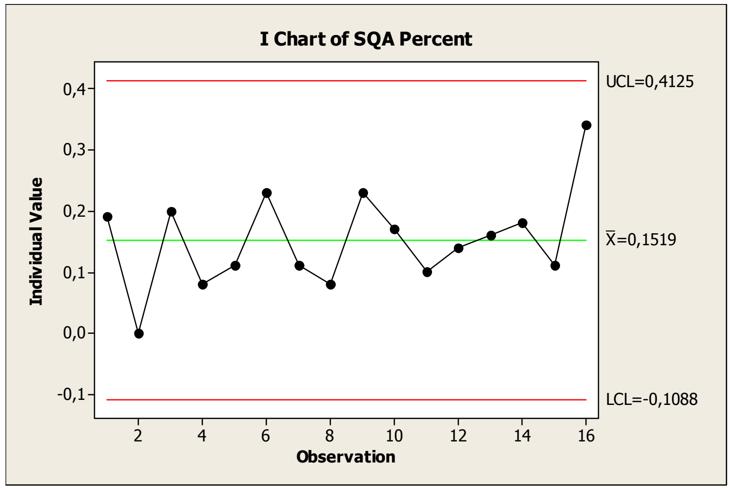


Figure E.61 Control Chart for Percent of Test Design Internal Review Effort of Test Development Cluster-2

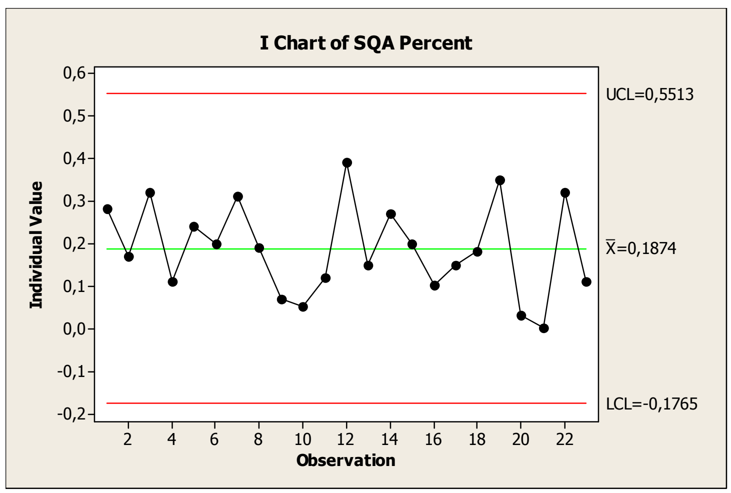


Figure E.62 Control Chart for Percent of Test Design Internal Review Effort of Test Development Cluster-3

Test Development Peer Review – Action Item Density

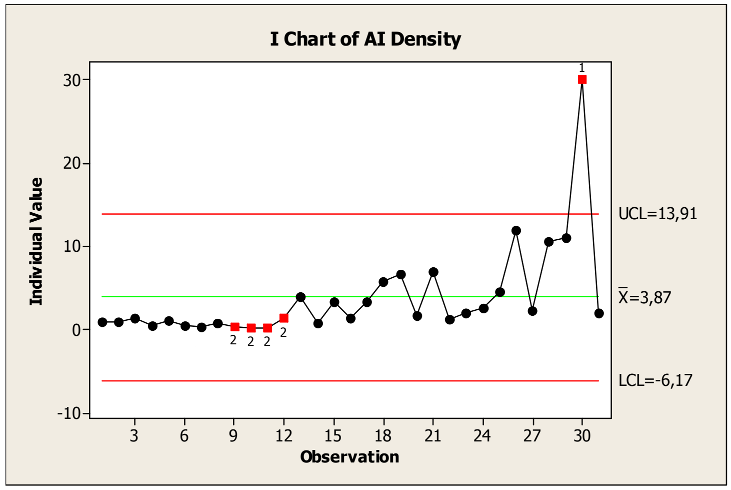


Figure E.63 Control Chart for Action Item Density of Test Development Peer Review Process (Overall)

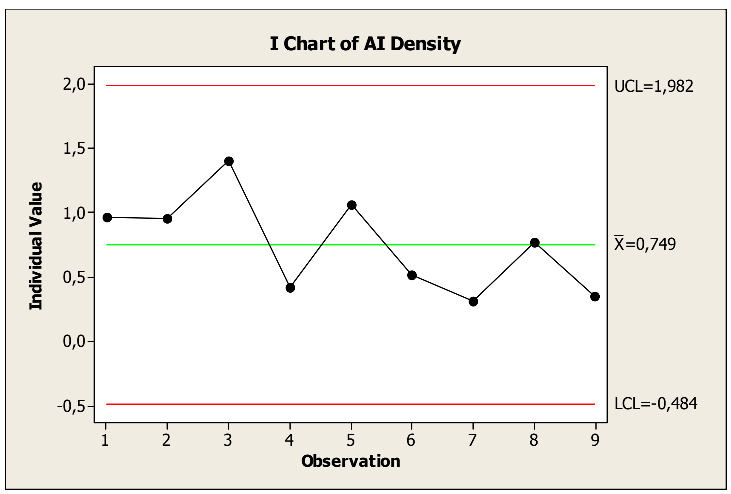


Figure E.64 Control Chart for Action Item Density of Test Development Peer Review Process Cluster-1

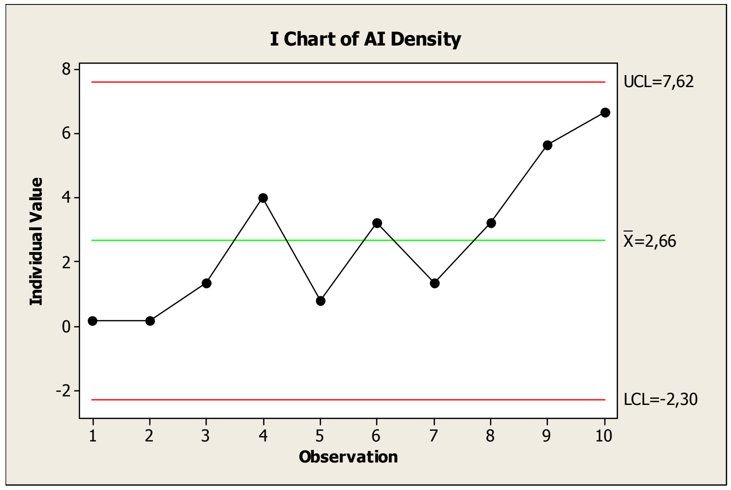


Figure E.65 Control Chart for Action Item Density of Test Development Peer Review Process Cluster-2

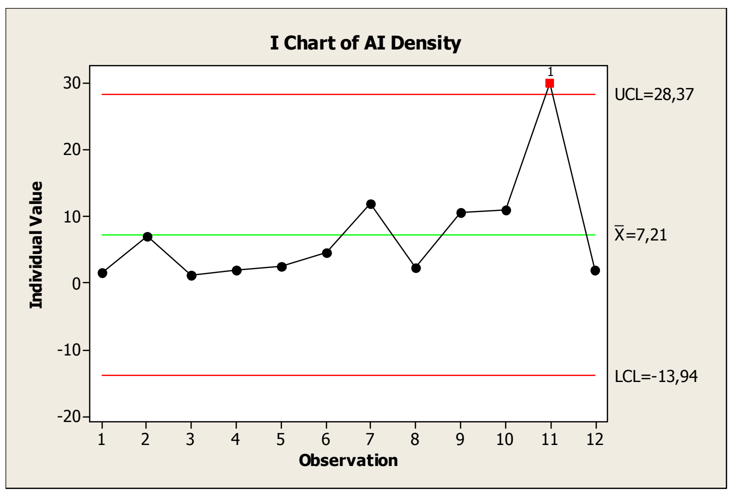


Figure E.66 Control Chart for Action Item Density of Test Development Peer Review Process Cluster-3

Test Development Peer Review – Action Item Detection Efficiency

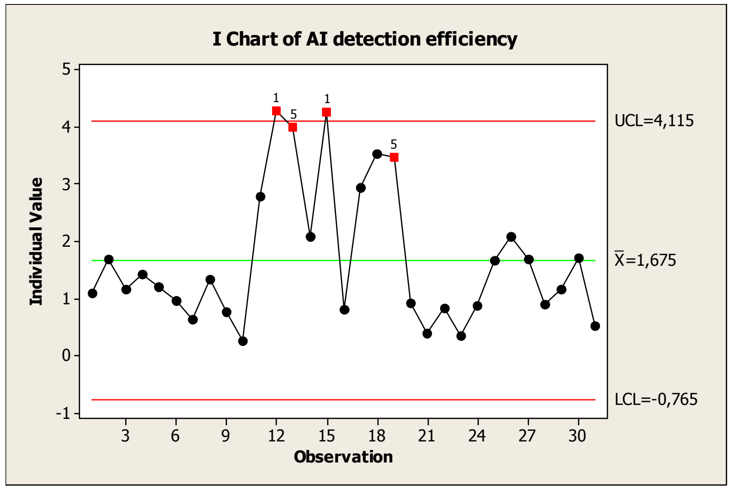


Figure E.67 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process (Overall)

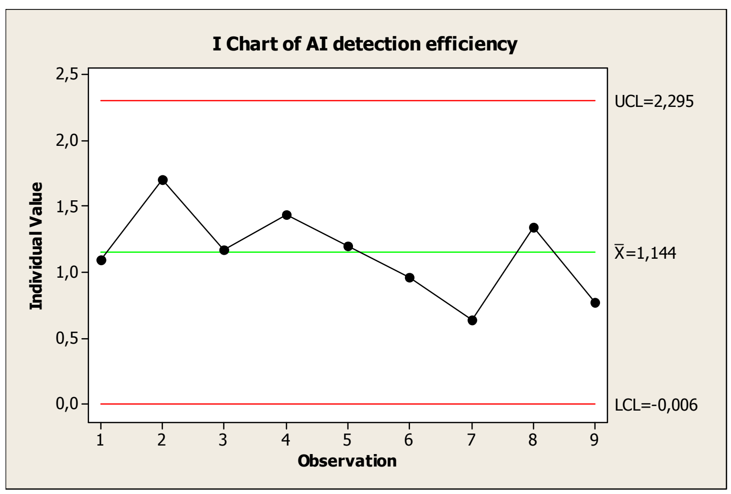


Figure E.68 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process Cluster-1

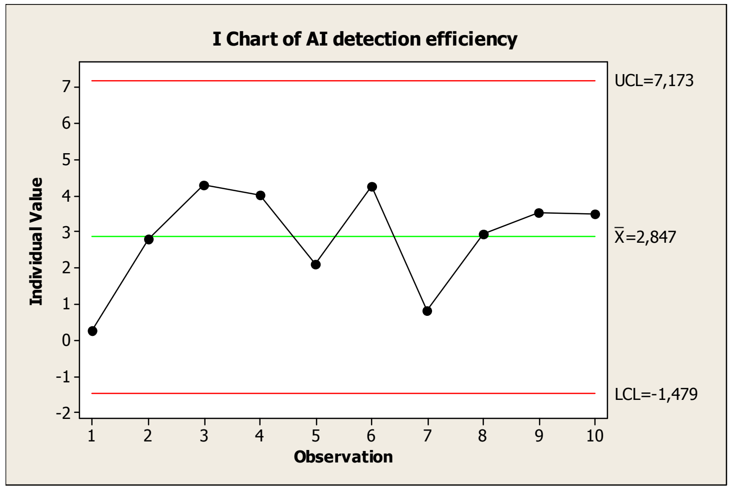


Figure E.69 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process Cluster-2

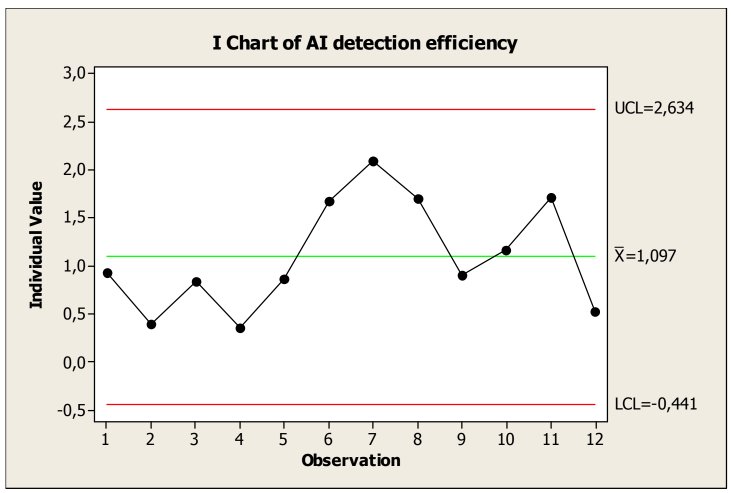


Figure E.70 Control Chart for Action Item Detection Efficiency of Test Development Peer Review Process Cluster-3

Test Development Peer Review – Action Item Resolution Efficiency

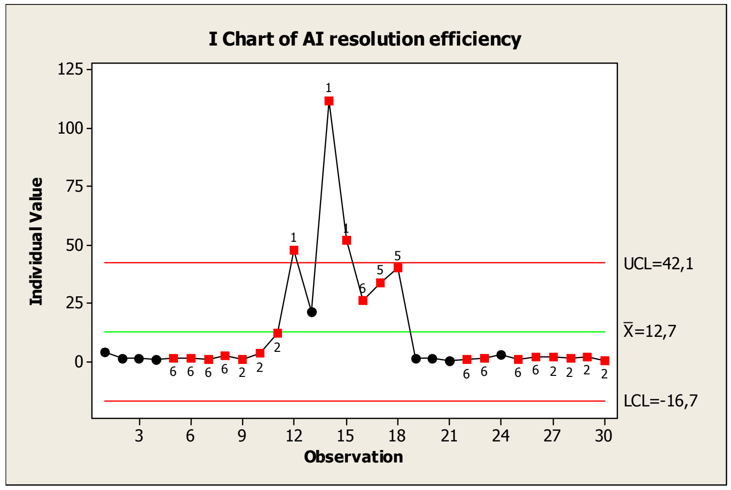


Figure E.71 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process (Overall)

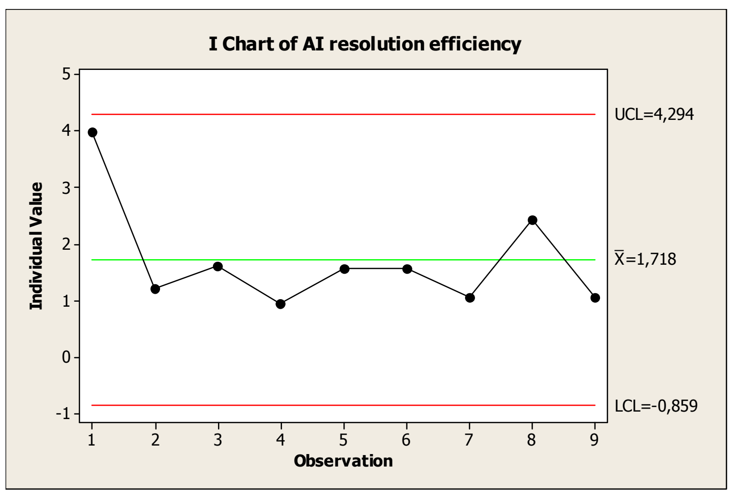


Figure E.72 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process Cluster-1

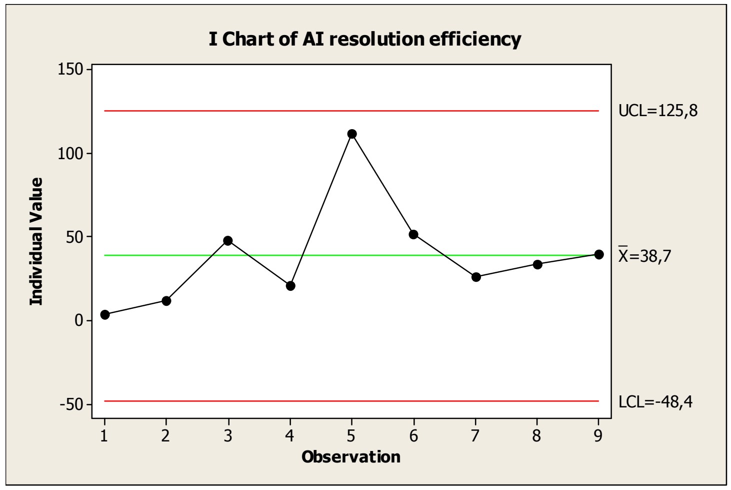


Figure E.73 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process Cluster-2

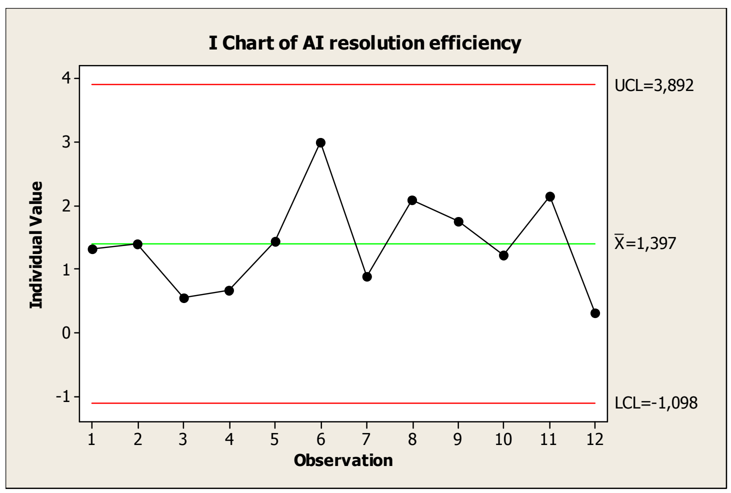


Figure E.74 Control Chart for Action Item Resolution Efficiency of Test Development Peer Review Process Cluster-3

Overall Reviews – Overall Review Effort per Test Case

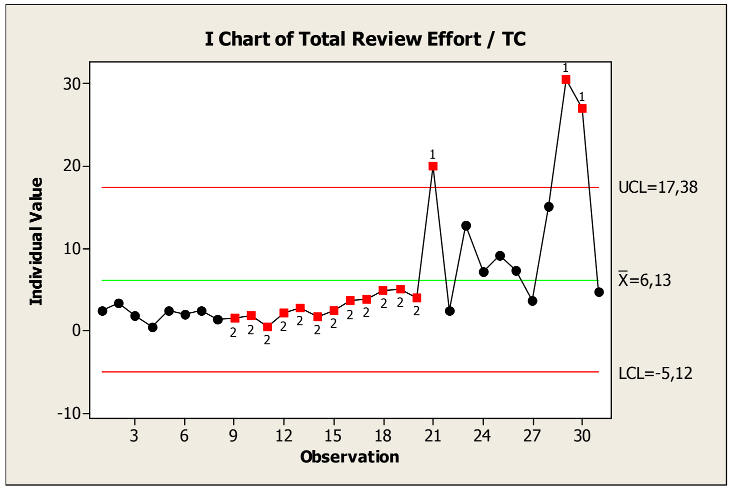


Figure E.75 Control Chart for Overall Review Effort per Test Case of Overall Reviews (Overall)

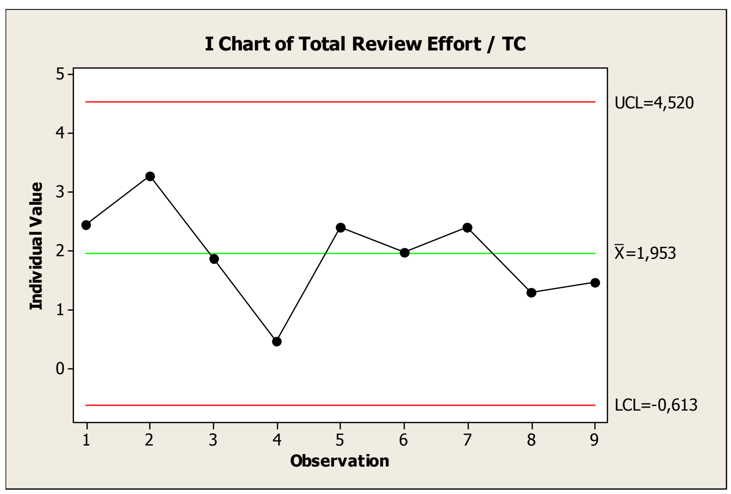


Figure E.76 Control Chart for Overall Review Effort per Test Case of Overall Reviews Cluster-1

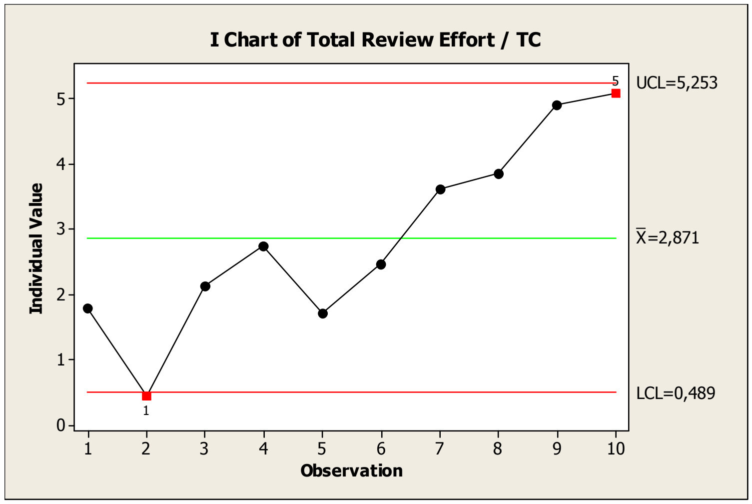


Figure E.77 Control Chart for Overall Review Effort per Test Case of Overall Reviews Cluster-2

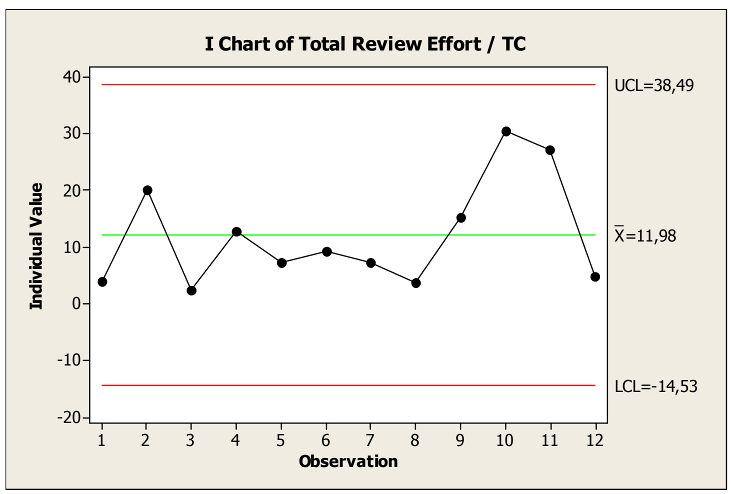


Figure E.78 Control Chart for Overall Review Effort per Test Case of Overall Reviews Cluster-3

CURRICULUM VITAE

Surname, Name: Tarhan, Ayça
Email: tarhan@ii.metu.edu.tr, ayca.tarhan@bg.com.tr

Education

MS in Computer Engineering, April 1999.
Dokuz Eylül University, Computer Engineering Department, İzmir.
Thesis Title : Commitment Management in Software Development.

BS in Computer Engineering, June 1995.
Ege University, Computer Engineering Department, İzmir.

High School, June 1990.
İzmir Science High School, İzmir.

Interested Areas

She has been working as instructor and consultant in the area of software engineering for 5 years. She pursues her studies with a focus on software quality, software measurement, and software engineering standards. She has expertise on SPICE assessments, personal software process, software acquisition, and object oriented technologies.

Experience

Bilgi Grubu Ltd.; Consultant/Instructor; March 2000-ongoing.

She has worked as team leader or consultant in projects involving software process assessment, software process improvement, software acquisition consultancy, and independent verification and validation.

She has participated in research projects supported by grants as a senior researcher: CORPUS Software Process Management System, ATRIUM Software Process Improvement Model, PERSONA Personal Software Engineering System.

Informatics Institute, Middle East Technical University, Ankara; System Analyst; 2001-2004.
Turkish Land Forces Command, two C4ISR Projects.

Informatics Institute, Middle East Technical University, Ankara; Part-time Instructor; 2002-2004.
MS Degree Courses offered (in English): Personal Software Process; Software Metrics; Object Oriented Analysis and Design.

Institute of Defense Sciences, Turkish Army Academy, Ankara; June 2002.
MS Degree Courses offered: Object Oriented Programming.

Computer Engineering Department, Dokuz Eylül University, İzmir; Research Assisstant; 1998-1999.

Computer and Informatics Department, Turkish Standards Institution, Ankara; System Analyst and Programmer; 1995-1997.

Certificates

Introduction to CMMI – Software Engineering Institute, 2005.
Personal Software Process – Software Engineering Institute, 2000.
SPICE Training Course – International SPICE Assessor Program, 1999.
Defining Software Processes Workshop – Software Engineering Institute, 1998.

Publications

"Investigating Suitability of Software Process and Metrics for Statistical Process Control", Springer Lecture Notes in Computer Science, EuroSPI Conference, Oct 2006, Joensuu, Finland (*to be presented*).

"Challenges of Acquisition Planning: Two Large System Acquisition Experiences", The 32th Euromicro Conference, Sep 2006, Cavtat/Dubrovnik, Croatia (*to be presented*).

"Pre-Contract Challenges: Two Large System Acquisition Experiences", Enterprise Integration, Idea Group Inc., Hershey-PA, 2006 (*book chapter - to be published*).

"Eliciting Information System Requirements From Business Process Models In The Pre-Contract Phase", SAVTEK 2006 Defense Technologies Congress, June 2006.

"Remarks from SPC Trial of an Emergent Organization", The 11th European SEPG Conference, June 2006, Amsterdam, Hollanda.

"Utilizing Business Process Models for Requirements Elicitation: A Large System Acquisition Experience", 29th Euromicro Conference, Sep 2003, Belek, Turkey.

"The Leap to Level 3: Can It be a Short Journey", SEI ESEPG Conference, June 2003, London, England.

"Managing Instructional Software Acquisition", Software Process Improvement and Practice Journal, Issue: 6, 2001.

"Tailoring ISO/IEC 12207 for Instructional Software Development", 26th Euromicro Conference, Sep 2000, Holland.

"A Distributed Environment for Managing Commitments in Software Development", ISCIS-XIV, Oct 1999, Kusadasi, Turkey.

"A Distributed Tool for Commitment Specification and Management", 25th Euromicro Conference, Sep 1999, Italy.

"Software Process Assessment Methods", Bilisim'98, Sep 1998, Istanbul, Turkey.