

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

THESIS PRESENTED TO
ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
PH.D.

BY
ZAROOUR, Mohammad

METHODS TO EVALUATE LIGHTWEIGHT SOFTWARE PROCESS ASSESSMENT
METHODS BASED ON EVALUATION THEORY AND ENGINEERING DESIGN
PRINCIPLES

MONTREAL, FEBRUARY 4, 2009

© Mohammad Zarour, 2009

THIS THESIS HAS BEEN EVALUATED
BY THE FOLLOWING BOARD OF EXAMINERS

Mr. Alain Abran, PhD, Thesis Supervisor
Department of Software Engineering and IT at the École de Technologie Supérieure

Mr. Jean-Marc Desharnais, PhD, Thesis Co-supervisor
Department of Software Engineering and IT at the École de Technologie Supérieure

Mr. Marcel Gabrea, PhD, President of the Board of Examiners
Department of Electrical Engineering at the École de Technologie Supérieure

Mr. Pierre Bourque, PhD, Examiner
Department of Software Engineering and IT at the École de Technologie Supérieure

Mr. Hakim Lounis, External Examiner
Computer Science Department, Université du Québec À Montréal

THIS THESIS WAS PRESENTED AND DEFENDED
BEFORE A BOARD OF EXAMINERS AND THE PUBLIC

JANUARY 28, 2009

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGMENTS

First and foremost I would like to extend my sincere gratitude to my supervisor Professor Dr. Alain Abran and my Co-supervisor Dr. Jean-Marc Desharnais. Without their patient guidance, this work would never have been started, carried out or finished nor would the road have been so pleasant. I learned a lot from their feedback through our continuous meetings. I was amazed by their insight and stamina. I would like to convey my warmest thanks to both Prof. Abran and Dr. Desharnais for investing their valuable time and resource on my behalf.

I would like to express my deep gratitude to my committee members for their time and effort in reviewing this work. I would also like to express my deep thanks to my GELOG colleagues and friends, with whom I have had the opportunity to enjoy and carry out this research.

Without industrial and academic interest, this research would not have been tested and improved. I would like to thank, at ETS, Dr. April for the time he took to conduct the evaluation of the S3^{mAssess} method and for the valuable comments which he made during the feedback session. I would like to thank Mr. Simon Alexandre in Belgium, for his evaluation of the Micro-evaluation method and for sharing his comments and suggestions. I would also like to thank Ms Wangenheim in Brazil, for her time conducting the evaluation of the MARES method and for sharing her comments and suggestions.

Finally, I would like to express my sincere thanks to my family for their loving support and understanding. My great mother has always believed in me and encouraged me to proceed with my studies. My wife Anaam, has encouraged me during the whole period of my study. She has shown an endless understanding of my research needs and has provided me with the means to concentrate on this research. My beloved children, Yman and Qusai, my source of joy and happiness, gave me the motivation to finalize this research. And I would also like to thank my brother and sisters for their support and encouragement.

MÉTHODES POUR ÉVALUER LES MÉTHODES D'ÉVALUATION DU PROCESSUS LOGICIEL LÉGER BÂTIES SUR LA THÉORIE DE L'ÉVALUATION ET LES CONCEPTS D'INGÉNIERIE DE CONCEPTION

ZAROOUR, Mohammad

RÉSUMÉ

Réaliser un processus de développement logiciel mature est devenu indispensable pour de nombreuses organisations de développement de logiciels. Un processus mature de développement de logiciels permet aux organisations de fournir à leurs clients des produits logiciels de haute qualité, livrés en temps et selon les budgets prévus.

Les organisations de développement de logiciels ont lutté pendant des décennies pour améliorer la qualité de leurs produits en améliorant leurs processus de développement logiciel. La conception d'un programme d'amélioration du processus de développement logiciel est exigeante et complexe. Un programme d'amélioration comprend deux processus principaux: le processus d'évaluation et le processus d'amélioration. Le succès du programme d'amélioration exige d'abord une évaluation réussie; le fait ne pas évaluer le processus de développement logiciel de l'organisation peut occasionner des résultats insatisfaisants.

L'évaluation des processus logiciels peut soit être utilisé pour déterminer la capacité d'une autre organisation, par exemple un sous-traitant, ou de déterminer et de comprendre le statut de l'actuel processus de l'organisation pour engager un processus d'amélioration. Le nombre croissant de processus d'évaluation disponibles, la norme ISO 15504 qui définit les exigences relatives à l'évaluation des processus et la popularité du modèle CMMI, illustrent la pertinence de l'évaluation des processus logiciels pour l'industrie du développement de logiciels.

Aujourd'hui, plusieurs méthodes sont disponibles pour évaluer la maturité et la capacité des processus de développement de logiciels. Ces méthodes sont basées sur des processus d'évaluation et d'amélioration de cadres de références bien connus, tels que le CMMI et ISO-15504. Le succès de ces méthodes d'évaluation et l'amélioration des cadres de référence est soutenu par des études post-développement sur la validité, la fiabilité et l'efficacité de ces méthodes. Malheureusement, de nombreux chercheurs constatent que de telles méthodes sont à trop grande échelle pour être mises en œuvre dans les TPE (Très Petites Entreprises). En conséquence, quelques chercheurs ont étudié le processus d'évaluation et d'amélioration dans les TPE et ont proposé aux organismes des méthodes d'évaluation généralement appelées "méthodes SPA légères", plus adaptée aux besoins de ces organisations TPE.

Les recherches actuelles dans le domaine des SPA mettent l'accent sur des propositions de méthodes d'évaluation faciles à utiliser, mais sans rechercher dans quelle mesure la conception de ces méthodes est liée à la vision de l'ingénierie de conception. Cet alignement imprécis avec la discipline de l'ingénierie soulève des questions quant à la pertinence et la représentativité des résultats obtenus par ces méthodes selon le point de vue de l'ingénierie.

En outre, bien que de nombreuses méthodes SPA actuellement disponibles offrent de l'aide et des orientations; elles n'adressent malheureusement que partiellement les éléments jugés essentiels au succès de la réalisation du SPA.

Cette thèse présente et discute l'évaluation de méthodes SPA. L'évaluation proposée dans cette thèse comprend deux volets: l'évaluation des méthodes SPA en utilisant des approches de conception descendante sur la base du point de vue de l'ingénierie ainsi qu'une approche de bas en haut pour évaluer le succès des méthodes de SPA. La théorie de l'évaluation des concepts est utilisée comme un cadre de référence pour développer formellement les deux méthodes d'évaluation.

Pour développer une première méthode d'évaluation en utilisant l'approche descendante, une étude exploratoire analytique des méthodes SPA à partir d'un point de vue de l'ingénierie de conception a été réalisée. La classification de Vincenti a été utilisée comme un instrument d'analyse. L'objectif de cette étude exploratoire est de placer les méthodes SPA développées dans le cadre de référence de l'ingénierie de conception et d'utiliser ce cadre de référence comme ligne directrice pour permettre de placer les nouvelles méthodes SPA à l'état de conception dans un même cadre référence en conception d'ingénierie. Pour développer la deuxième méthode d'évaluation en utilisant l'approche de bas en haut, un examen systématique de la littérature a été réalisée pour extraire l'ensemble des éléments nécessaires au succès des méthodes de SPA fondées sur les exigences, les observations, les leçons apprises et les recommandations qui ont été expérimentées dans l'industrie et publiées dans des livres, des conférences et des revues.

Le processus de développement des deux méthodes d'évaluation a ensuite été vérifié en utilisant un ensemble de critères de vérification. Ensuite, la proposition des méthodes d'évaluation a été testée par la réalisation de trois études de cas. La première méthode d'évaluation serait surtout utile pour les concepteurs de la nouvelle méthode de SPA au cours de la phase de conception, tandis que la seconde méthode d'évaluation serait utile pour les concepteurs et les praticiens des méthodes SPA pour vérifier le succès de la méthode d'évaluation en question.

Ce projet de recherche constitue un point d'entrée pour étudier l'alignement des méthodes du SPA de conception avec les principes de conception de l'ingénierie. Ce projet de recherche fait aussi la lumière sur la réalisation réussie des résultats d'évaluation en étudiant le succès des éléments qui doivent être supportées par des méthodes d'évaluation séparé des processus d'amélioration. Les méthodes d'évaluation proposées dans cette thèse ont de grands avantages pour les méthodes de SPA conçues avant tout pour les TPE parce que ces méthodes d'évaluations, contrairement aux méthodes plus connues, ne sont pas encore appuyées par des études complètes sur leur fiabilité et leur efficacité.

Mots-clés : évaluation de processus de logiciel, amélioration de processus de logiciel, qualité de logiciel, théorie d'évaluation, ingénierie de la conception.

METHODS TO EVALUATE LIGHTWEIGHT SOFTWARE PROCESS ASSESSMENT METHODS BASED ON EVALUATION THEORY AND ENGINEERING DESIGN CONCEPTS

ZAROOUR, Mohammad

ABSTRACT

Achieving a mature software development process has become essential for many software organizations. A mature development process permits software organizations to provide their customers with a high quality software product delivered on time and within budget.

Software organizations have been struggling for decades to improve the quality of their products by improving their software development processes. Designing an improvement program for a software development process is a demanding and complex task. This task consists of two main processes: the assessment process and the improvement process. A successful improvement process requires first a successful assessment; failing to assess the organization's software development process could create unsatisfactory results.

Software processes assessment (SPA) can be used either to determine the capability of another organization, for subcontracting purposes, or to determine and understand the status of the organization's current processes to initiate an improvement process. The increasing number of assessment approaches available; the ISO 15504 standard that sets out the requirements for process assessment; and the popularity of the CMMI model, illustrate the relevance of software process assessment for the software development industry.

Currently, several methods are available to assess the maturity and capability of a software development process based on well-known software process assessment and improvement frameworks such as CMMI and ISO-15504. The success of these assessment methods and improvement frameworks is supported by post-development studies on the validity, reliability and effectiveness of these methods. Unfortunately, many researchers consider that such methods are too large to implement in SME organizations. As a result, many researchers have studied process assessment and improvement in SME organizations and proposed assessment methods, usually called lightweight SPA methods, suitable to the organizations' needs.

The current research in the SPA field focuses on proposing convenient and easy-to-use assessment methods, without investigating to what extent the design of these methods is related to the engineering design perspective. This unclear alignment with the engineering discipline raises questions about the relevance and representativeness of the results produced by these methods from an engineering viewpoint. Moreover, although numerous SPA methods are currently available which offer help and guidance, unfortunately they only partially address evidences found essential for achieving an SPA success.

This thesis presents and discusses the evaluation of lightweight SPA methods. The evaluation is two-fold: evaluating the SPA methods design using a top-down approach and based on engineering viewpoints and evaluating the success of SPA methods using a bottom-up approach. The evaluation theory concepts are used as a framework to formally develop both evaluation methods.

To develop the first evaluation method using the top-down approach, an exploratory analytical study of SPA methods from an engineering design viewpoint has been conducted. Vincenti's classification has been used as a tool for this analysis. The aim of this exploratory study is to put the developed SPA methods into an engineering design framework, and use this framework as a guideline to put the new SPA methods to be designed into the same engineering design framework. To develop the second evaluation method using the bottom-up approach, a systematic literature review was conducted to extract the set of evidences for the success of the SPA method based on requirements, observations, lessons learned and recommendations which have been formulated within the industry and published in books, conferences and journals.

The development process of the two evaluation methods has then been verified using a set of verification criteria and the proposed evaluation methods were tested by conducting three case studies. The first evaluation method would be useful mainly for the designers of new SPA methods during the design phase, while the second evaluation method would be useful for both designers and practitioners of SPA methods to verify the success of the assessment method in question.

This research project forms an entry point to study the alignment of SPA methods design with engineering design principles and sheds light on achieving successful assessment results by studying the successful evidences that should be supported by assessment methods separated from the improvement process. The proposed evaluation methods in this thesis have great benefits for SPA methods designed mainly for SME organizations, because these assessments methods, contrary to well-known methods, lack comprehensive studies on their reliability and effectiveness.

Keywords: software process assessment SPA, software process improvement SPI, software quality, evaluation theory, engineering design.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
 CHAPTER 1 SOFTWATRE PROCESS IMPROVEMENT	 11
1.1 Introduction.....	11
1.2 Well-known SPI frameworks.....	13
1.2.1 Capability maturity model – CMM.....	13
1.2.2 Capability maturity model integrated - CMMI	16
1.2.3 ISO 15504	16
1.3 Success factors for SPI programs based on well-known SPI frameworks	17
 CHAPTER 2 SOFTWATRE PROCESS ASSESSMENT.....	 19
2.1 Introduction.....	19
2.2 Assessment Methods related to the well-known SPI frameworks.....	19
2.2.1 SCAMPI.....	20
2.2.2 ISO 15504 assessments.....	20
2.3 Problems specific to SME organizations	21
2.3.1 Financial problems.....	22
2.3.2 Undefined organizational structure and responsibilities.....	22
2.3.3 Organizational success is based on individual skills	23
2.3.4 Long term return on investment (ROI)	23
2.4 Lightweight assessment methods.....	23
2.4.1 MARES.....	23
2.4.2 FAME – Fraunhofer IESE assessment method	24
2.4.3 TOPS – toward organized process in SMEs	25
2.4.4 RAPID – Rapid assessment for process improvement for software development.....	26
2.4.5 Micro-Evaluation assessment	27
2.4.6 Express process appraisal (EPA) assessment methods.....	28
2.4.7 Software process improvement initiation framework - SPINI	29
2.4.8 A Modular software process Mini-Assessment method	29
2.4.9 S ^{3m} Mini-Assessment method - S ^{3mAssess}	30
2.5 Summary	30
 CHAPTER 3 RESEARCH OBJECTIVES AND METHODOLOGY	 34
3.1 Introduction.....	34
3.2 Research motivation.....	35
3.3 Research goal	36
3.4 Research objectives.....	37

3.5	Research scope.....	37
3.6	Research methodology.....	38
CHAPTER 4 A FRAMEWORK TO COMPARE LIGHTWEIGHT SOFTWARE PROCESS ASSESSMENT METHODS		43
4.1	Introduction.....	43
4.2	Comparison methods in detail	45
4.2.1	Halvorsen's taxonomy	45
4.2.2	Anacleto comparison method	46
4.3	Proposed comparison framework	48
4.3.1	Characteristics description.....	50
4.3.2	Application of the comparison framework	51
CHAPTER 5 REFERENCE DISCIPLINES.....		54
5.1	Evaluation theory concepts.....	54
5.1.1	Evaluation method classifications.....	55
5.1.2	Basic components of an evaluation.....	57
5.2	Overview of the evaluation components in the SPA context	58
5.2.1	Target	59
5.2.2	Criteria	60
5.2.3	Yardstick	62
5.2.4	Data gathering techniques.....	63
5.2.5	Synthesis techniques	64
5.2.6	Evaluation process	64
5.3	Different technological viewpoints.....	67
5.3.1	Shaw's viewpoint.....	67
5.3.2	Simon's viewpoint	68
5.3.3	Vincenti's viewpoint.....	69
5.3.4	Abran et al. viewpoint.....	70
CHAPTER 6 DEVELOPING EVALUATION METHODS BASED ON ENGINEERING DESIGN PRINCIPLES TO EVALUATE SPA METHODS		72
6.1	Introduction.....	72
6.2	The target	72
6.3	The evaluation criteria	72
6.3.1	Mapping Vincenti's classifications to the SPA methods design context.....	73
6.3.1.1	Fundamental design principles evaluation criteria	75
6.3.1.2	Criteria and specifications evaluation criteria.....	80
6.3.1.3	Theoretical tools evaluation criteria.....	81
6.3.1.4	Quantitative data evaluation criteria	82
6.3.1.5	Practical considerations evaluation criteria	83
6.3.1.6	Instrumentalities evaluation criteria.....	83

6.4	The yardstick.....	85
6.5	Data gathering technique	86
6.6	The synthesis technique	87
6.7	The evaluation process.....	87

CHAPTER 7 DEVELOPING EVIDENCE-BASED EVALUATION METHODS TO EVALUATE LIGHTWEIGHT SPA METHODS		89
7.1	Introduction.....	89
7.2	The systematic literature review	90
7.2.1	Planning the review.....	91
7.2.1.1	Identify the needs for a systematic review.....	91
7.2.1.2	Develop the review protocol.....	92
7.2.1.3	Review protocol.....	94
7.2.2	Conducting the review	94
7.2.2.1	Identification of research	94
7.2.2.2	Selection of primary studies.....	95
7.2.2.3	Study quality assessment	95
7.2.2.4	Data extraction.....	95
7.2.2.5	Data analysis.....	96
7.2.2.6	Reporting the review.....	101
7.3	Detailed discussion of SPA success evidences.....	101
7.3.1	ISO/IEC 15504-3 (ISO/IEC 2003): guidance on performing an assessment.....	102
7.3.2	SCAMPI requirements (CMU/SEI 2006).....	104
7.3.3	Watts Humphrey (Humphrey and Kitson 1987).....	105
7.3.4	Van Han Loon (Loon, Cass et al. 2004)	106
7.3.5	Christiane Gresse Wangenheim (Wangenheim, Varkoi et al. 2006).....	107
7.3.6	Toshihiro Komiyama (Komiyama, Sunazuka et al. 2001)	108
7.3.7	Fabbrini (Fabbrini, Fantini et al. 2003).....	110
7.3.8	Literature review summary	111
7.4	Evaluation components definitions.....	117
7.4.1	The target	117
7.4.2	The evaluation criteria	118
7.4.2.1	SPA method evaluation criteria (SPA-MEC)	119
7.4.2.2	SPA tool evaluation criteria (SPA-TEC)	120
7.4.2.3	SPA procedure evaluation criteria (SPA-PEC).....	120
7.4.2.4	SPA documentation evaluation criteria (SPA-DEC)	121
7.4.2.5	SPA user evaluation criteria (SPA-UEC)	121
7.4.3	The yardstick.....	121
7.4.3.1	Yardstick based on SPA method.....	122
7.4.3.2	Yardstick based on SPA support tool	125
7.4.3.3	Yardstick based on SPA procedure.....	127
7.4.3.4	Yardstick based on SPA documentation.....	129
7.4.3.5	Yardstick based on SPA users	131

7.4.4	Data gathering technique	133
7.4.5	The synthesis technique	133
7.4.6	The evaluation process.....	134
CHAPTER 8 RESEARCH APPROACH VERIFICATION PROCESS.....		135
8.1	Introduction.....	135
8.2	Type of research conducted	135
8.3	Verification techniques	136
8.4	Validation based on grounded theory	139
CHAPTER 9 CONDUCTED CASE STUDIES		142
9.1	CASE STUDY 1 – EVALUATION OF THE S3 ^{mAssess} ASSESMENT METHOD	142
9.1.1	Software maintenance process	142
9.1.1.1	Software maintenance maturity model S3 ^M	143
9.1.1.2	The software maintenance assessment method S ^{3mAssess}	144
9.1.2	Evaluation procedure	146
9.1.3	S3 ^{mAssess} assessment method evaluation results based on the bottom-up evidence-based approach	147
9.1.3.1	Strength points	148
9.1.3.2	Weakness points.....	149
9.1.4	S ^{3mAssess} method evaluation results based on Top-down approach	152
9.1.5	Results of the feedback session.....	1566
9.1.5.1	Quick feedback results	156
9.1.5.2	Detailed feedback results	1577
9.2	CASE STUDY 2 – EVALUATION OF THE ‘MICRO-EVALUATION’ METHOD	162
9.2.1	Introduction to the ‘Micro-Evaluation’ method.....	162
9.2.2	Evaluation procedure	163
9.2.3	Evaluation of the Micro-Evaluation method based on the Bottom-up approach.....	164
9.2.3.1	Strength points	165
9.2.3.2	Weakness points.....	166
9.2.4	Micro-Evaluation assessment method evaluation results based on Top-Down evaluation method	170
9.3	CASE STUDY 3 – EVALUATION OF THE MARES ASSESMENT METHOD	173
9.3.1	Introduction.....	173
9.3.2	Evaluation procedure	174
9.3.3	MARES assessment method evaluation results based on Bottom-up evaluation method.....	175
9.3.3.1	Strength points:	176
9.3.3.2	Weakness points.....	178

9.3.4	MARES assessment method evaluation results based on Top-Down evaluation method	180
9.4	Comparisons among the three SPA methods.....	182
9.4.1	Comparison based on the Bottom-up evaluation method	182
9.4.2	Comparison based on the Top-up evaluation method.....	184
CONCLUSION.....		187
ANNEX I	THE EVALUATION TOOL DEVELOPED FOR THE EVALUATION METHOD BASED ON ENGINEERING DESIGN PRINCIPLES	191
ANNEX II	SUMMARY OF THE PUBLICATIONS REVIEWED IN THIS THESIS AND FOUND TO DISCUSS SPECIFIC SUCCESS EVIDENCES RELATED TO THE SPA METHODS.....	194
ANNEX III	REQUIREMENTS FOR CLASS-C CMMI APPRAISAL METHODS	200
ANNEX IV	SUMMARY OF THE REQUIREMENTS FOR PERFORMING ASSESSMENTS AS STATED IN ISO 15504-3	208
ANNEX V	SUMMARY OF THE EVIDENCES FOUND NECESSARY FOR THE SUCCESS OF THE ASSESSMENT METHODS AND THEIR REFERENCES	213
ANNEX VI	THE EVALUATION TOOL DEVELOPED FOR THE EVIDENCE BASED EVALUATION METHOD	237
BIBLIOGRAPHY		240

LIST OF TABLES

	Page
Table 2.1	Requirements of CMMI appraisal method classes21
Table 2.2	Comparison of the principles and processes of lightweight assessment methods32
Table 4.1	Categorization of characteristics in Halvorsen taxonomy46
Table 4.2	Descriptions of some examples for excluded characteristics48
Table 4.3	Categorization of characteristics in the proposed comparison framework49
Table 4.4	The comparison results of the seven lightweight SPA methods.....52
Table 5.1	Evaluation methods proposed by Worthen55
Table 5.2	Evaluation methods proposed by House.....56
Table 6.1	Loon's Assessment procedure summary79
Table 6.2	Summary of the SPA method yardstick - Top-Down approach85
Table 6.3	Sample of the Evaluation Tool - Top-Down approach.....87
Table 7.1	Archiving format for the set of publications and their evidences94
Table 7.2	Sample of the documents discussing SPA success evidences102
Table 7.3	Publications relating the five classes of collected success evidences112
Table 7.4	The collected set of evidences and their frequencies grouped into five main classes115
Table 7.5	Statistics summary of the success evidences found in the literature117
Table 7.6	SPA method yardstick (SPA-MYS).....122
Table 7.7	SPA supportive tool yardstick (SPA-TYS).....125
Table 7.8	SPA procedure yardstick (SPA-PYS).....127
Table 7.9	SPA documentation yardstick (SPA-DYS)129
Table 7.10	SPA user yardstick (UYS)131

Table 7.11	Sample of the evaluation tool - Bottom-Up approach	134
Table 9.1	Bottom-up evaluation results of $S^{3mAssess}$ method	148
Table 9.2	Number of questions per engineering design class	1533
Table 9.3	Top-down evaluation results of $S^{3mAssess}$ method	154
Table 9.4	Modified Bottom-up evaluation results of $S^{3mAssess}$ method - based on feedback session	159
Table 9.5	Modified Top-down evaluation of $S^{3mAssess}$ method – based on feedback session	161
Table 9.6	Bottom-up evaluation method results of the Micro-Evaluation method .	164
Table 9.7	Top-down evaluation method results of Micro-Evaluation method	172
Table 9.8	Bottom-up evaluation method results of MARES assessment method ...	175
Table 9.9	Top-down evaluation method results of MARES method	181

LISTE OF FIGURES

	Page
Figure 1.1	Generic software process improvement model.....11
Figure 1.2	Generic environment of software process development and process improvement Program.14
Figure 1.3	The five levels of software process maturity.15
Figure 1.4	A summary of the CSFs as stated by the literature and interviews.18
Figure 3.1	Summary of the reference disciplines.....38
Figure 3.2	Research methodology.....42
Figure 5.1	Components of an evaluation.....57
Figure 5.2	Interrelations between components.....58
Figure 5.3	Breakdown graph for criteria elicitation methods.61
Figure 5.4	Main phases of the evaluation process.....65
Figure 5.5	Detailed activities for the proposed evaluation process phases.67
Figure 6.1	Steps to build Top-Down evaluation method.73
Figure 6.2	Vincenti's classifications breakdown graph – evaluation criteria.75
Figure 7.1	Bottom-Up evaluation method development process.90
Figure 7.2	Main phases of a systematic literature review.91
Figure 7.3	Data model for the conducted literature review.....93
Figure 7.4	Main classes of the collected evidences.....96
Figure 7.5	Extending the developed Bottom-Up evaluation method by adding a third level114
Figure 7.6	Research community interests in the different classes of success evidences.....118
Figure 7.7	SPA process evaluation criteria tree.119

Figure 9.1	S3m process model.	145
Figure 9.2	Strengths comparisons between the three SPA methods.	183
Figure 9.3	Weaknesses comparisons among the three SPA.....	183
Figure 9.4	Comparison of strength contribution among the three SPA methods.....	184
Figure 9.5	Comparison among satisfied design issues of the three SPA methods based on the Top-down approach	185
Figure 9.6	Comparison among unsatisfied design issues of the three SPA methods based on the Top-down approach.....	186

LISTE OF ABBREVIATIONS

CMM	Capability Maturity Model
CMMI	Capability Maturity Model Integration
SPA	Software Process Assessment
SPI	Software Process Improvement
VSE	Very Small Enterprises
SME	Small and Medium Size Enterprises
SCAMPI	Standard CMMI Appraisal Method for Process Improvement
SPICE	Software Process Improvement and Capability dEtermination
SPM	Software Process Matrix
GQM	Goal Question Metrics
SQM	Software Quality Management
QIP	Quality Improvement Paradigm
SWEBOK	Software Engineering Body of Knowledge

INTRODUCTION

The software industry has become an important economic activity in industrialized countries. Investments in this field are measured in billions of dollars. In parallel, the number of software organizations has increased, varying in size, types of products produced and the development process used. (Humphrey 1989) stated that “The mature development processes have great impact on the quality of the software product”. For software organizations to provide their customers with quality software products, the software development process needs to be improved. The impact of improvement methodologies on software quality has been studied by (Kuilboer and Ashrafi 2000; Ashrafi 2003), and the findings reveal that each SPI framework has a different level of impact on software quality factors.

Moreover, the direct relationship between product quality and the adopted development process, as well as the need of assessment for continuous improvement, is stated in (ISO/IEC 1992): “Product quality is highly dependent on processes used in its creation and the way toward product quality is to have available and to utilise a proven, consistent, reliable method for software process assessment and use the results in a continuous improvement programme”. (Kim and Yang 2006) also stated that “There has always been recognition that process assessment can be a strong and effective driver for process improvement. Much empirical evidence has accumulated demonstrating the benefits that can be derived from an assessment-based software process improvement”.

Consequently, achieving a mature software development process has become a strategic core competency in many organizations (Komi-Servio 2004). A mature development process permits software organizations to provide their customers with quality software products delivered on time and within budget, as stated in (M. C. Paulk et al. 1995): “The primary intended long-term benefit of high process maturity is high-quality software meeting customer requirements, delivered on time and within budget”. (Humphrey 1992) also stated that “The key objective of a mature software process is to produce quality products to meet

customer's needs". Hence, many software organizations are struggling to improve their development process capabilities.

Today, several methods are available to assess the maturity and capability of a software development process based on well-known software process assessment and improvement frameworks such as CMMI (CMMI Product Team 2006) and ISO-15504 (ISO/IEC 2003-2006). The success of these assessment methods and improvement frameworks is supported by studies on their validity, reliability and effectiveness; see for example (El-Emam and Goldenson 1995; Goldenson, Emam et al. 1997; Lee, Jung et al. 2001; Jung and Goldenson 2002).

One should note that capability is not the same as performance. Software process performance is the actual result achieved from following certain software processes, while software process capability is what is expected from the process. Results achieved (performance) differ from results expected (capability).

Software process improvement approaches can be classified into two main categories (Gorschek 2006):

1. Top-Down approach (Prescriptive).

The top-down approach compares an organization's process with a generally accepted standard process. Process improvement is then the elimination of differences between an existing process and a standard process. The assumption is that, once the process is changed, the generated product will be improved or at least the risk in generating new software will be reduced (Thomas and McGarry 1994). Examples of this category are the model-based approaches such as CMMI and ISO 15504.

2. Bottom-up approach (Inductive).

"The bottom-up approach assumes that process change be driven by an organization's goal, characteristic, product attributes, and experiences. Change is defined by a local domain instead of universal set of accepted practices" (Thomas and McGarry 1994).

“This approach assumes that every development organization must first completely understand its process, products, software characteristics, and goals before it can select a set of changes meant to improve its process” (Thomas and McGarry 1994). An examples of this category is the six sigma (**6σ**) and Quality Improvement Paradigm (QIP) (Basili 1985).

Why conduct an assessment?

“Understanding how to implement SPI successfully is arguably the most challenging issue facing the SPI field” (Dyba 2000). Successful implementation of SPI requires successful assessment first. For instance, a successful SPA process should show weaknesses in the development process and should recommend realistic actions to overcome one of the barriers for SPI success as stated in (Dyba 2000) which is SPA recommendations that are too ambitious.

Showing up weaknesses and recommending realistic actions would help software organisations to establish improvement priorities before attempting a solution (Humphrey 1989). Hence, software process assessment is conducted to (Zahran 1998):

1. Understand and determine the organisation’s current software engineering practices, and to learn how the organisation works.
2. Identify strengths, major weaknesses and key areas for software process improvement.
3. Facilitate the initiation of process improvement activities and involve opinion leaders in the change process.
4. Provide a framework for process improvement actions.

Research issues

Although many successful SPI case studies have been published (Bhandari, Halliday et al. 1993; Herbsleb, Carleton et al. 1994; McGarry, Pajerski et al. 1994; Goldenson and Herbsleb 1995; Diaz and Sligo 1997), criticism of current SPI approaches also exist (Bollinger and

McGowan 1991; Gray and Smith 1998). Criticising the current SPI trends would focus the researchers' attention on the weaknesses of the current research and would also open the door for new research trends.

Other issues in the SPI field are still being raised and more research is needed to address them:

1. The first issue is related to the efficiency and effectiveness of the current SPI approaches. Much of the research into this issue of the SPI approaches has been performed by consultants involved in their promotion (Card 2004). To provide industry with objective and rigorous studies related to this issue, academic researchers should conduct more research into the efficiency and effectiveness of SPI approaches.
2. The second issue is related to repeatability and reproducibility of software process assessments. When conducting the same assessment with different groups, would the results be the same? Similarly, if an assessment with one group is conducted at different intervals, would the results be the same? What is the variation between the different assessment results and what is the acceptable range of such variations? These and other related questions should be addressed thoroughly.
3. The third issue is related to the validity of a software process. Viewing software development as a process has helped in identifying the different dimensions of software development and the problems that need to be addressed in order to establish effective practices (Fuggetta 2000). Accordingly, several research initiatives have investigated different areas related to the software model, including software process modeling and support, process improvement and related measures and empirical studies. Unfortunately, and as a result of the complex nature of the software process, the literature on software process related areas has not yet provided a full understanding of the nature of the software process. Such an understanding is vital before pursuing further research to evaluate the software process. The need for such understanding has

been highlighted by (Gray and Smith 1998): “Full understanding of the nature of the software process should be in place before and to underpin the design of process evaluation schemes. It is the case that process assessment methods are here before the required understanding”.

Despite the amount of results produced so far in increasing the quality and effectiveness of software development processes, software process research is undergoing a crisis that is visible through a number of symptoms (Fuggetta 2000):

1. Most technologies developed by the software process community have not been transferred into industrial use.
2. The number of papers on software process modeling and technology presented at conferences and published in journals is decreasing.

There is an increasing feeling that the community is unable to produce innovative and effective contributions: (Fuggetta 2000) concluded that these crises in the software process emphasize the need to rethink the adopted approach in studying and supporting software processes.

4. The fourth issue is related to the data acquisition mechanisms used in software process assessment methods. SPA methods employ questionnaires to conduct the assessment. Unfortunately, questionnaires are problematic in the sense that questions can have semantics with different interpretations by different participants. One of the consequences of this problem is the impact on the repeatability of the assessment results as discussed in the second issue. Ambiguity and inconsistency in the interpretation of questions is currently the norm in practice augmenting the subjectivity on the part of auditors (Gray and Smith 1998) .
5. The fifth issue is related to assessment-based SPI in very small, small and medium size organizations. The well-known SPI frameworks, as well as their related assessment

methods, have been seen by many researchers as being too complicated to implement and require too much effort and cost (Mäkinen, Varkoi et al. 2000; Rout, Tuffley et al. 2000; Habra, Renault et al. 2002; Wangenheim, Anacleto et al. 2004; CETIC 2006; FUNDP-CETIC – Software Quality Lab 2006; Wangenheim, Varkoi et al. 2006; McCaffery, Taylor et al. 2007; Pettersson, Ivarsson et al. 2008). As a result, many researchers have studied process assessment and improvement in SME and VSE organizations and proposed assessment methods suitable for such organization's needs; such assessment methods are presented in detail in chapter 2.

Consequently, the research issues mentioned previously for the SPI in general are not only valid and raised again here for SPI in SMEs and VSE organizations, but are more challenging for these organizations due to their organizational and cultural nature which may differ from that of other types of organizations. Moreover, SPA methods designed to fit the needs of SME and VSE organizations lack studies on their validity, reliability, and effectiveness as in the SPA method of CMMI and ISO 15504, for example.

Organization of this thesis

This thesis contains eleven chapters (including the introduction and the conclusion) and six Annexes. The current chapter outlines research issues, organization of the thesis and the definitions used.

Chapter one presents an overview of the field of software process improvement and of the different well-known SPI frameworks currently in use. This chapter also discusses the potential benefits of software process improvement and the barriers facing the implementation of the improvement process in SME organizations.

Chapter two presents a general overview of software process assessment methods and gives a classification to these methods based on the organization's environment suitability for the implementation of the assessment methods. A description of each method for each class is presented in more detail.

Chapter three presents the research goal, objectives and research methodology. The research methodology is divided into three phases; the details of each phase are presented in this chapter.

Chapter four presents the different approaches found in the literature to compare different assessment methods and discuss their benefits and drawbacks. A detailed framework to compare lightweight assessment methods is then presented.

Chapter five presents the referenced disciplines that will be used as a foundation for this research, namely evaluation theory concepts and engineering design principles.

Chapter six presents the development of the Top-down evaluation method; this method is dedicated to the evaluation of the assessment methods design process. The method uses evaluation theory components as a framework to build the proposed evaluation method. This chapter also presents the engineering design principles used as a basis to build the proposed evaluation method.

Chapter seven presents the development of the Bottom-up evaluation method. The method also uses the evaluation theory components as a framework to build the proposed evaluation method. This chapter also presents the success evidences used as a basis to build the proposed evaluation method.

Chapter eight presents the verification process used to verify the research approach adopted in this thesis and defines the verification criteria on which the verification process is to be conducted.

Chapter nine presents the three case studies carried out to evaluate three SPA methods; the first case study discusses the evaluation of the S3^{mAssess} assessment method which has been designed for the evaluation of the software maintenance process. The evaluation was conducted by the author of this method in Canada. The second case study discusses the

evaluation of the OWPL Micro-Evaluation assessment method. The evaluation was conducted by the authors of this method in Belgium. The third case study discusses the evaluation of the MARES assessment method. The evaluation was conducted by the authors of this method in Brazil.

The Conclusion chapter summarizes the results of this thesis, its contributions and limitations and suggests future work.

Finally, this thesis contains six ANNEXs. Annex I presents the evaluation tool developed for the Top-down evaluation method based on engineering design viewpoint. Annex II presents a summary of the publications that discussed success evidences related to the SPA methods. Annex III presents CMMI appraisal requirements for method of class C. Annex IV presents the requirements of performing assessment as stated in ISO 15504-3. Annex V presents the frequency of success evidences found in the literature. Annex VI presents the evaluation tool developed for the bottom-up, evidence based evaluation method.

Definitions

Before discussing software process assessment and improvement in detail, it is necessary to understand the vocabulary used in describing the software process; one should note that the models considered in this thesis are far from being limited to a mathematical formula – actually, a ‘model’ is a description of how to conduct the process of software development.

The key terms used in this thesis are defined next. These terms are based on the definitions in the ISO 15504 part 1 standard (ISO/IEC 2004). Adopting these standard definitions is necessary to avoid the non uniform definitions of the terms adopted by different researchers.

Process: Set of interrelated or interacting activities which transforms inputs into outputs.

Process assessment: A disciplined evaluation of an organizational unit’s processes against a ‘Process Assessment Model’.

Process Assessment Model: A model suitable for the purpose of assessing process capability, based on one or more Process Reference Models.

Process capability determination: A systematic assessment and analysis of selected processes within an organization against a target capability, carried out with the aim of identifying the strengths, weaknesses and risks associated with deploying the processes to meet a particular specified requirement.

Process improvement: Actions taken to change an organization's processes so that they more effectively and/or efficiently meet the organization's business goals.

Process improvement programme: All the strategies, policies, goals, responsibilities and activities concerned with the achievement of specified improvement goals.

Process improvement project: Any subset of the process improvement programme that forms a coherent set of actions to achieve a specific improvement.

Process improvement sponsor: The individual or entity, internal or external to the organizational unit being assessed, who requires the process improvement to be performed, and provides financial or other resources to carry it out.

Process reference model: A model comprising definitions of processes in a life cycle described in terms of process purpose and outcomes, together with an architecture describing the relationships between the processes.

Assessment sponsor: The individual or entity, internal or external to the organizational unit being assessed, who requires the assessment to be performed, and provides financial or other resources to carry it out.

Assessment team: One or more individuals who jointly perform a process assessment.

Assessor: An individual who participates in the rating of process attributes.

Assessment participant: An individual who has responsibilities within the scope of the assessment. Examples include but are not limited to the assessment sponsor, assessors, and organizational unit members.

CHAPTER 1

SOFTWARE PROCESS IMPROVEMENT

1.1 Introduction

The research work in software process was initiated in the 1980's (Rozman, Horvat et al. 1994); it had been recognized by researchers and practitioners that many management and support activities are needed to produce effective software organizations - for example: cost and schedule problems as well as inadequate product quality are an outcome of inadequate management. This inadequate management can be due to inadequate time and resources allocated to project managers.

Most of SPI frameworks are based on a general principle of four fairly straightforward stages, as shown in Figure 1.1:

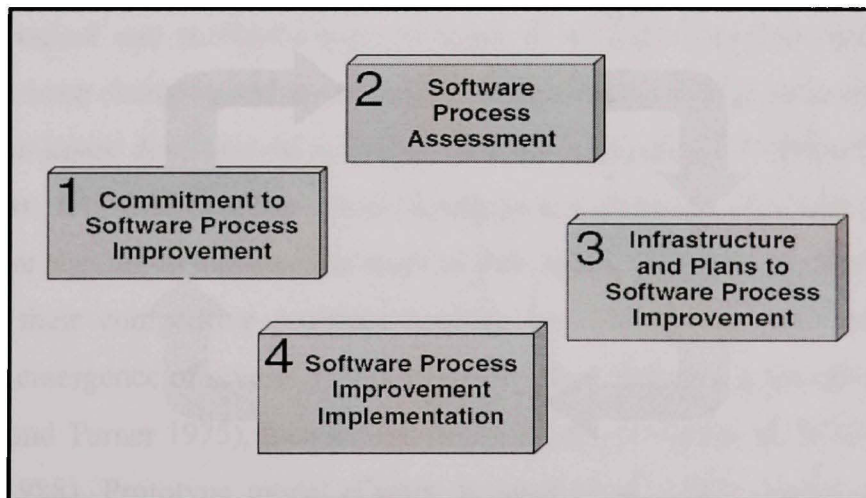


Figure 1.1 Generic software process improvement model.
(Villalón, Cuevas et al. 2002)

- Stage 1: Commitment to improvement. Its objective is to obtain the support of senior management to carry out the improvement project;

- Stage 2: Software process assessment. Its objective is to obtain strengths and weaknesses of the process assessed with respect to a software process model;
- Stage 3: Infrastructure and Action Plan. Its objective is to provide the needed infrastructure to carry out improvement in selected processes, and to create the plan to follow in order to define and implement improvement in these selected processes;
- Stage 4: SPI implementation. Its objectives are to define the new processes elected following the previous plan, and to implement them in pilot projects. Finally, improvement must be institutionalized.

As a result of the widespread of software either as a stand alone product or as part of embedded systems, the amount and complexity of software has increased. In addition, market requirements have led to tightened lead-time requirement, i.e. the software should be produced more and more rapidly. Unfortunately, this can cause a great amount of quality problems (Komi-Servio 2004).

Software process improvement is, usually, conducted in a complex environment. In addition to the technological and methodological changes in software development, the business strategies have been changing and are changing (Komi-Servio 2004); software development has started as in-house development activities, then development of Commercial off-the-shelf COTS products, followed by open source development approach. Through these changes, companies were seeking to enhance the ways of developing software, to improve quality, or to strengthen their competitive position. Seeking to enhance the ways of development resulted in the emergence of several software development models e.g. iterative enhancement model (Basili and Turner 1975), Incremental model (Mills, O'Neill et al. 1980), Evolutionary model (Gilb 1988), Prototype model (Curtis, Krasner et al. 1987), Spiral model (Boehm 1988), V-model (GMOD 1992) and the Agile methods (Abrahamson, Salo et al. 2002). Seeking to improve quality resulted in the emergence of several software process improvement models e.g. CMMI. The purpose of improvement is to enhance the software development process in order to raise software quality (Basili and Caldiera 1995). Software quality is evaluated by the software customers. The customers along with the competition in

the market affect the cost, functionality and quality requirements. SPI program includes both assessment and improvement activities; the assessment process aims at discovering the strengths and weaknesses in the organization's development process, and then produces recommendations used to build the action plan to improve the software development process. Both assessment and improvement processes require certain infrastructure and resources to be available in the organization to ensure a successful improvement program. Once the SPI process is terminated, a reassessment process is conducted again to measure the improvements achieved and the whole process is repeated again until the required improvements are totally achieved. As a summary, Figure 1.2 illustrates the complex environment for the assessment and improvement process.

The two main organizations that initiated the work in software process assessment and improvement were the ISO organization and the Software Engineering Institute (SEI) at Carnegie Melon University (CMU) in USA.

1.2 Well-known SPI frameworks

There are several model-based frameworks that can be used to assess the current maturity level of an organization's software processes. In the following sub-sections the commonly used SPI frameworks are presented.

1.2.1 Capability maturity model – CMM

CMM has been developed by the SEI Institute at Carnegie Melon University. It is centered on standardizing the contents of processes according to a predefined number of practices. The practices of CMM are organized into Key Process Areas (KPA). Each KPA is placed within one of five levels of process maturity.

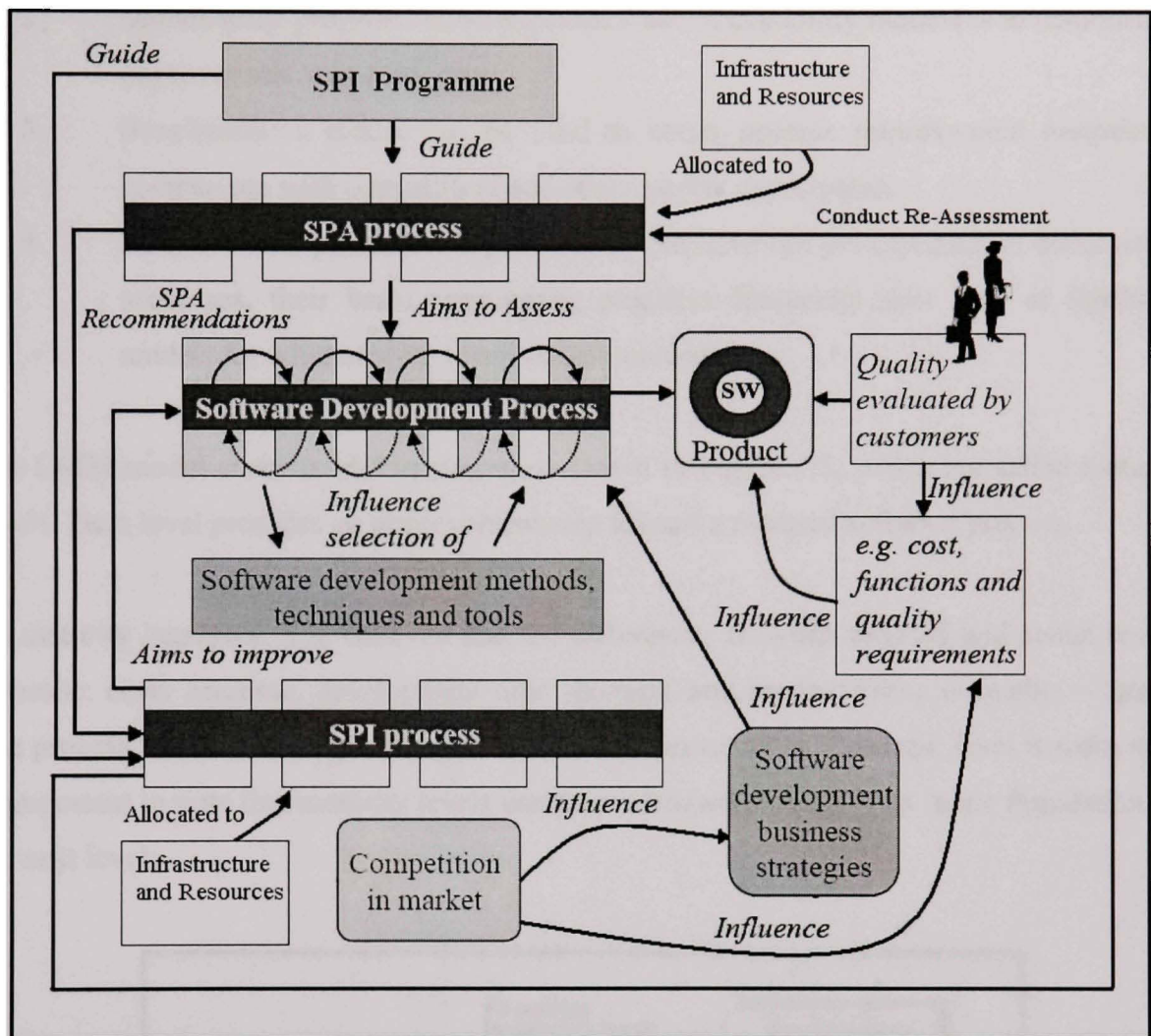


Figure 1.2 Generic environment of software process development and process improvement Program.

Adapted from (Komi-Sevio 2004)

The overall purpose of capability models is to establish a process improvement roadmap upon which a route can be drawn from “where we are today” to “where we want to be”. The capability models define the characteristics of good processes and avoid prescribing how the processes must be performed. CMM is used to (Cepeda, Staley et al. 2004):

1. Verify process content: capability models encapsulate basic industry knowledge for an organization to use to help improve quality, customer satisfaction, productivity and cycle time.

2. Demonstrate progress: another primary use of capability models is to demonstrate improvement year over year.
3. Benchmark: a model can be used to verify process improvement progress in comparison with competitors and other market participants.
4. Structure new processes: organizations that have not yet captured, in documented processes, their basic engineering practices frequently will look at capability models for a list of what needs to be included.

The CMM model consists of five stages, as shown in Figure 1.3, which are called maturity levels. Each level provides an improvement step toward a matured software process.

As maturity increases, it is believed that the differences between targeted and actual results decrease; costs decrease, development time shortens; and productivity and quality increase. The process becomes more predictable as rework is prevented and the risk level is reduced. It is important to note that maturity levels cannot be skipped as each level is the foundation for the next level.

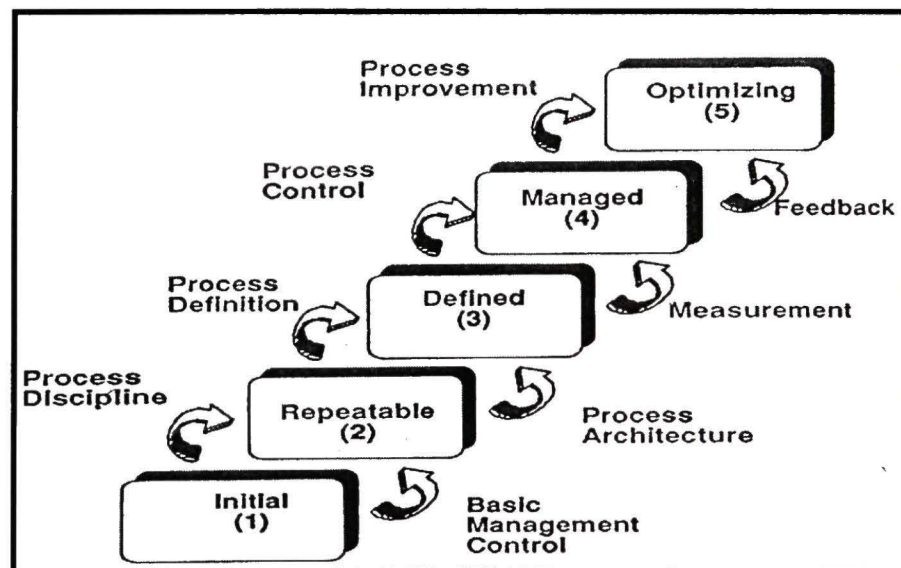


Figure 1.3 The five levels of software process maturity.
(Humphrey 1992)

1.2.2 Capability maturity model integrated - CMMI

The CMM framework is divided into different versions, each of which being related to a specific discipline, e.g. CMM for software SW-CMM, CMM for System engineering SE-CMM and Integrated product development CMM IPD-CMM (Ahren, Clouse et al. 2004).

In the late 1990's, the SEI began an effort to integrate all of the approaches to CMM into a common integrated version - CMMI, which was initially released in 2000. At the same time, the SEI has attempted to generalize CMMI so that it can be used to evaluate any organization's ability to manage processes.

CMMI comes in two basic versions: Staged and Continuous Representation. Both are based on the same process areas, but they are represented differently and thus address SPI in different ways. Staged Representation is aimed towards assessing and improving overall organizational maturity based on the maturity levels. Continuous Representation is adapted towards assessing and improving individual process areas based on the organizations' needs. CMMI Continuous Representation can be used to address the specific process areas as relevant to a specific organization.

1.2.3 ISO 15504

ISO 15504 provides a structured approach for the assessment of processes. Process assessment is based on a two-dimensional model composed of a process and a capability dimension. The use of process assessment within an organization should encourage (ISO/IEC 2003-2006):

1. The culture of continuous improvement and the establishment of proper mechanisms to support and maintain this culture.
2. The engineering of processes to meet business requirements.
3. The optimization of resources.

This encouragement leads the organization to become a capable organization that maximizes its responsiveness to customer and market requirements, minimizes the full life-cycle costs of its products and as a result maximize end-user satisfaction.

The ISO 15504 series (ISO/IEC 2003-2006) consists of five documents:

- 15504-1 - Concepts and Vocabulary;
- 15504-2 - Performing an Assessment;
- 15504-3 - Guidance on Performing an Assessment;
- 15504-4 - Guidance on Use for Process Improvement and Process Capability Determination;
- 15504-5 - An Exemplar Process Assessment Model.

Mapping between ISO 15504 and CMMI is relatively easy since they have been developed in parallel with observers from each in their development process: they have been aligned ‘on purpose’. Some differences between the two models still remain and include: a) process areas that are present in one, but not in the other, b) ISO 15504 has only the Continuous Representation (not Staged).

1.3 Success factors for SPI programs based on well-known SPI frameworks

An extensive empirical study and literature survey of the Critical Success Factors (CSF) that impact SPI implementation has been discussed in (Rainer and Hall 2003) and (Niazi, Wilson et al. 2006), where the experiences, opinions and views of practitioners have been analyzed in order to identify factors that have a positive impact on the implementation of SPI programs.

Focusing on these factors, as suggested in (Niazi, Wilson et al. 2006), offers SPI practitioners short-term opportunities for implementing practices that have impact on the SPI

implementation process. The collected set of critical factors of both literature and empirical study is shown in Figure 1.4.

As can be seen from Figure 1.4, most of the critical success factors deal with management issues that affect the success of the improvement process. Although there are some common factors for both assessment and improvement process such as commitment, involvement and assigning responsibilities, these critical success factors did not touch the specific factors that affect the assessment process.

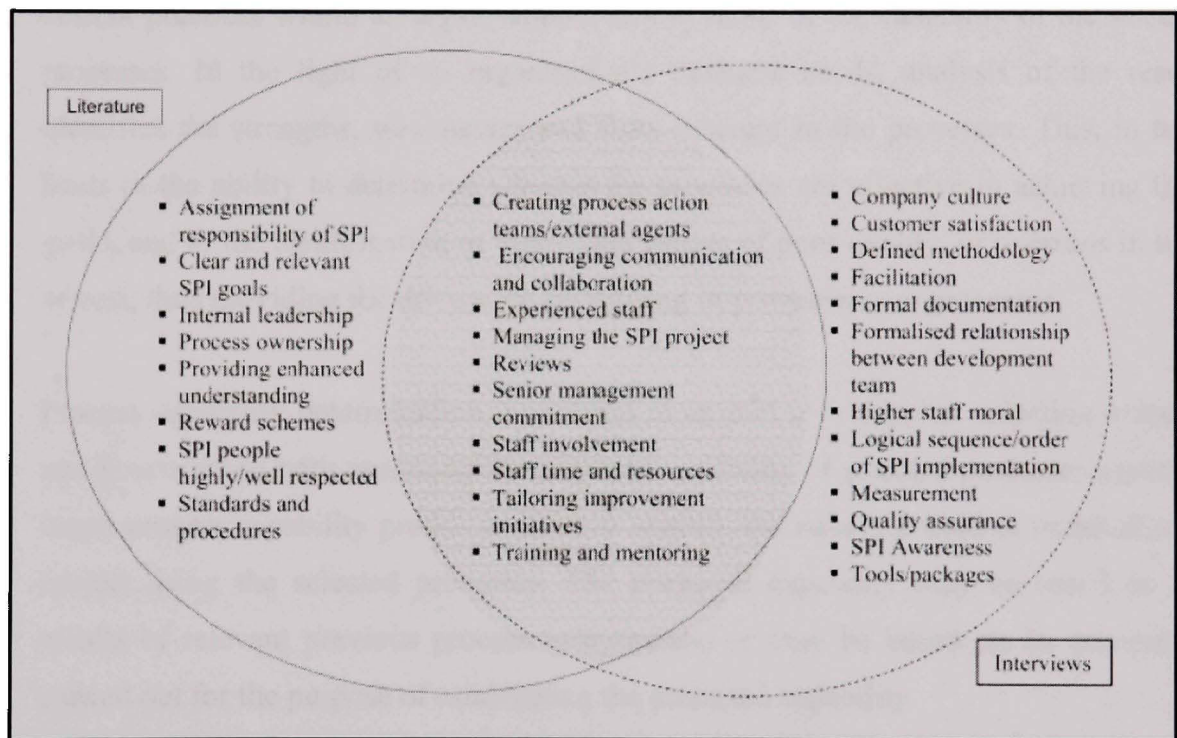


Figure 1.4 A summary of the CSFs as stated by the literature and interviews.
(Niazi, Wilson et al. 2006)

CHAPTER 2

SOFTWARE PROCESS ASSESSMENT

2.1 Introduction

Process assessment has two principal contexts for use (ISO/IEC 2004):

1. Process improvement context, process assessment provides means of characterizing the current practices within an organisational unit in terms of the capability of the selected processes. In the light of an organisation's business needs, analysis of the results identifies the strengths, weaknesses and risks inherent in the processes. This, in turn, leads to the ability to determine whether the processes are effective in achieving their goals, and to the identification of significant causes of poor quality, or overruns in time or cost, thus providing the drivers for prioritising improvements to processes.
2. Process capability determination is undertaken as part of a supplier selection exercise and is concerned with analyzing the proposed capability of selected processes against a target process capability profile in order to identify the risks involved in undertaking a project using the selected processes. The proposed capability may be based on the results of relevant previous process assessments, or may be based on an assessment carried out for the purpose of establishing the proposed capability.

2.2 Assessment Methods related to the well-known SPI frameworks

The assessment methods related to the well-known SPI frameworks, discussed in section 1.2, are SCAMPI and the ISO 15504 conformant assessment method; these methods are presented briefly hereafter.

2.2.1 SCAMPI

The CMMI appraisal methodology is based on several appraisal requirements called Appraisal Requirements for CMMI or ARC. These requirements are a basis on which appraisals can be developed. The official implemented appraisal method for CMMI is called Standard CMMI Appraisal Method for Process Improvement - SCAMPI. This method was developed to meet all the requirements described in ARC as well as to be compliant to ISO 15504 (Ahren, Clouse et al. 2004). In general, CMMI supports three classes of appraisals (CMMI-SUT 2006):

- **Class A:**
A Full and comprehensive method covering the entire CMMI model and providing a maturity level of the organization as a whole. SCAMPI (V1.2) is a Class A assessment method;
- **Class B:**
A less in depth class than class A; this class concentrates on areas that need attention and gives no overall maturity rating and is considered as beneficial as an initial assessment method;
- **Class C:**
This class is often called “a quick look” at specific risk areas.

The requirements of the CMMI appraisal method classes as stated in (CMMI-SUT 2006) are summarized in Table 2.1.

2.2.2 ISO 15504 assessments

ISO/IEC 15504 defines requirements for performing an assessment and provides an exemplar assessment model that meets such requirements. Another example of assessment models is SPICE for SPACE of S4S (Cass, Volcker et al. 2000; Volcker and Cass 2000; Volcker and Cass 2001) used by the European Space Agency.

As a result of having similar requirements, ISO 15504 assessments conducted to initiate an SPI program are very similar to SCAMPI (CMMI Class A).

Table 2.1

Requirements of CMMI appraisal method classes
(CMMI-SUT 2006)

Requirements	Class A	Class B	Class C
Types of objective evidence gathered	Documents and interviews	Documents and interviews	Documents or interviews
Ratings generated	Goal ratings required	Not allowed	Not allowed
Organizational unit coverage	Required	Not required	Not required
Minimum team size	4	2	1
Appraisal team leader requirements	Lead appraiser	Person trained and experienced	Person trained and experienced

2.3 Problems specific to SME organizations

Small and Medium (SME) size organizations are encountering a number of problems when trying to adopt an SPI approach, especially those approaches designed for large organizations. The main characteristics of this type of organizations are as follows (Varkoi and Makinen 2000):

1. Flexible and fast to adopt new cultures.
2. Highly concentrated to their main business i.e. they have only a few products.
3. Lean organisations (some even without any administrative staff).
4. Extensively customer-oriented.
5. Undertaking small projects (typically 6-18 person months).
6. Networked with other small companies.
7. Dependant on skilful individuals.
8. Overloaded with work but enthusiastic.
9. With limited resources for business development or process improvement.

The main problems facing SME organizations are treated in the following paragraphs:

2.3.1 Financial problems

Usually SME organizations are fighting to survive, and strive to provide the client with a working version of the software without the funding to pay enough attention to the software quality and documentation processes. Hence, small companies tend to be less focused on the process through which the software is written. The main reason for this unplanned software development process is that small organizations do not have enough financial support to cover the high cost of adopting an SPI approach. Adopting such an SPI approach is usually a costly process that also requires a lot of time, raising another problem which is the cost of time, i.e. small organizations suffer from the lack of financial support which is a major factor in improving their processes. Kautz (Kautz 1998) has found that the external financial support is one of four critical success factors for SPI in small organizations.

2.3.2 Undefined organizational structure and responsibilities

As an organization becomes larger, the need for Software Quality Management - SQM becomes vital. SQM requires the creation of specialized teams or groups with specific responsibilities for designing, coding and testing. By doing so, the organization would have control over the software development process.

Usually, there is a greater tendency that large organizations and large software development projects will utilize SQM; the reason might be that larger organizations have more resources and therefore better possibilities for engaging in SQM (Kautz and Ramzan 2001). Although SQM is as vital for small organizations as it is for large organizations, small organizations do not seem to utilize SQM as regularly as large organizations (Kautz and Ramzan 2001); there are no specialized teams in small organizations where the organizational structure is undefined and wide responsibilities are assigned to persons involved in the software development life cycle, as mentioned in (Ward, Laitinen et al. 2000): “The organizational

chart, if one even exists, is delineated in broad, intuitive strokes and is composed of roles and responsibilities defined more by this week's crisis than by any reified notion of corporate structure".

2.3.3 Organizational success is based on individual skills

Usually, the lack of success of small organizations is related to talented people not working to a standardized development process which is very risky for the survival of the organizations. Knauber (Knauber, Muthig et al. 2000) has discussed the importance of these talented people in small organizations: "We found that one major factor to consider is the influence of a few key individuals— perhaps the company founder or those who play key roles because of their very strong skills". Ward (Ward, Laitinen et al. 2000) has also indicated this critical factor when he said: "Small companies live and die on the engineering talent they are able to hire".

2.3.4 Long term return on investment (ROI)

Richardson (Richardson 2002) has stated that small organizations cannot make large investments in SPI and long term ROI, and suggests that having a fast ROI is one of the requirements in any SPI method designed for small organizations.

2.4 Lightweight assessment methods

Lightweight assessment methods have been created to overcome the problems facing SMEs organizations which are starting SPI initiatives.

2.4.1 MARES

A Methodology for Software Process Assessment in Small Software Companies – the MARES model, has been built by researchers from the UNIVALI University and CenPRA

research center in Brazil. MARES was designed to support process improvement in the context of small software organizations; considering their specific characteristics and limitations, this MARES model is built in conformity to ISO 15504 (ISO/IEC 2003-2006). MARES enhanced the process assessment model mainly by integrating a context-process model in order to support the selection of relevant processes and a process-risk model to support the identification of potential risks and improvement suggestions. The MARES assessment method is divided into five main parts:

1. Planning:

In this phase, the assessment is organized and planned; at the end of this phase the resulted assessment plan is revised and documented.

2. Contextualization:

In this phase, the organization is characterized in order to understand its goals, products and software process. Questionnaires and interviews are used as a means to collect data.

3. Execution:

The selected processes are assessed in detail.

4. Monitoring and control:

All activities during the assessment are monitored and controlled. Corrective actions are initiated if necessary, and the plan is updated accordingly.

5. Post-mortem:

Once the assessment is finished, a brief post-mortem session is held by the assessors to discuss and evaluate the performance of the assessment.

2.4.2 FAME – Fraunhofer IESE assessment method

FAME (Beitz, Emam et al. 1999) is a stand-alone assessment method which is based on the ISO 15504 (ISO/IEC 1998) assessment method. The use of the FAME method has the following benefits (Hamann, Beitz et al. 2001):

- Allows for the performance of either a SPICE, currently known as ISO 15504 standard, or a BOOTSTRAP Assessment;
- Focuses on relevant business processes to guide process improvement efforts;
- Provides a cost-efficient and reliable method to show a better return on investment for the improvement program;
- Provides a tailorable approach for performing assessments;
- Provides an approach that allows an organisation to compare its results with similar businesses based upon ISO 15504;
- Provides a method that is applicable for small to large organizations.

FAME contains supplementary added value elements that have been developed through practical experiences from the worldwide ISO 15504 trials and from Fraunhofer IESE research results. These added value elements are the business Focus, efficiency, reliability and benchmarking.

2.4.3 TOPS – toward organized process in SMEs

TOPS is a rapid software process assessment method created by Florence University in Italy, to promote innovation in IT small and medium enterprises. The TOPS method tries to find a compromise between accurate results and low costs; this method is based on a two part questionnaire (Cignoni 1999):

- The first part is made by phone and is organized into five sections that collect general data about the enterprise, define company characteristics and information regarding regional industry survey to evaluate the knowledge of European initiatives to support enterprises in their SPI experiments;
- The second part is the basis for the assessment and is compiled during a meeting with the enterprise. This part is organized into three sections: collecting general data about the software development unit; assessing the organizational and technological characteristics of the software development unit; and the assessment of software

processes with respect to three specific processes including requirement analysis, verification and tests, and joint review. These processes are based on the ISO 15504 (ISO/IEC 1998).

This approach restricted its assessment to three processes for time constraints, to provide more time for discussion. This method also avoids “the risk of difficult topics, as for instance configuration management, that may need explanation diverting the focus of meeting.” (Cignoni 1999).

2.4.4 RAPID – Rapid assessment for process improvement for software development

RAPID, developed by the Software Quality Institute at Griffith University (Queensland, Australia), defines an approach to assessment that delivers consistent evaluations of process capability based upon an intensive investigation of the operations of the organization. The approach is based upon the following principles (Rout, Tuffley et al. 2000):

1. The assessment is conducted within a one-day timeframe.
2. The assessment is based upon an assessment model of limited scope, with a standard set of eight processes.
3. The competence and experience of the assessors is seen as of primary importance. A team of two assessors with experience in performing full-bodied assessments based upon ISO 15504. (ISO/IEC 1998) is used for a RAPID assessment.
4. Data collection is limited to the single technique of moderated discussions by performers of the processes, the management team and other members of the organization.
5. Generation of ratings of capability is performed by a process of consensus-gathering involving all of the participants in the discussion, rather than by the judgment of the assessors. Restricting the assessment to a single day rather than a more intense three to four day assessment enables small organizations to participate in a process capability assessment. Most organizations are willing to invest a day of their time and resources.

The RAPID method employs a defined assessment model of restricted scope based upon, and compatible with, the Process Reference Model of ISO 15504-2. The model includes eight processes, including:

1. Requirements Gathering.
2. Software Development.
3. Project Management.
4. Configuration Management.
5. Quality Assurance.
6. Problem Resolution.
7. Risk Management.
8. Process Establishment.

2.4.5 Micro-Evaluation assessment

The Micro-Evaluation assessment method is based on the OWPL assessment model (CETIC 2006). The OWPL model has been developed based on the ISO 15504 model. The OWPL approach for software process assessment and improvement is known as the OWPL gradual framework. This gradual framework involves a series of gradual assessments (Alexandre, Renault et al. 2006): a micro-evaluation, an OWPL-evaluation and a SPICE or CMM assessment. The nested assessments can be used either separately or in successive stages in the SPI process.

Micro-evaluation assessment is the first step in the gradual framework and is a simplified model designed to reduce costs as much as possible and to give a first look at the assessed organization. The purpose of this model is:

- To make the assessed SME organizations aware of weaknesses as well as potential expected improvements;

- To determine the priorities of subsequent stages of evaluation and improvement procedures.

This method is based on an interview with a person having sufficient knowledge of organizational activities through a questionnaire. The questionnaire covers six key processes, called axes, selected as the most pertinent and most important to the target organization. These axes are (Habra, Renault et al. 2002):

1. Quality assurance.
2. Customer's management.
3. Subcontractor's management.
4. Project management.
5. Product management.
6. Training and human resources management.

Although the Micro-Evaluation framework was initially developed for evaluating IT organizations in the Wallonia area in Belgium, a few other similar regions with a lot of similarities to Wallonia area have applied this framework. For example, in the province of Quebec in 2004, a research project was initiated at ETS called Amélioration de la Performance des Petites Entreprises Québécoises – APPEQ. The aim of this project was to help SMEs organizations improve their quality, productivity and performance.

2.4.6 Express process appraisal (EPA) assessment methods

The Express Process Appraisal - EPA method (McCaffery, McFall et al. 2005; Wilkie, McCaffery et al. 2007) was developed by the Centre for Software Process Technologies in the UK to assess software processes within SME organisations that have little or no experience in software process improvement programs. This method is conformant with the ARC 1.1 requirements for a CMMI class-C method. The designers of this EPA method selected the six most appropriate process areas (to software companies within Northern

Ireland) at CMMI maturity level 2 since the justification for starting a process improvement exercise with these process areas was already well established, being present at the first level in the model. The following process areas were selected: Requirements Management; Configuration Management; Project Planning; Project Monitoring and Control; Measurement and Analysis; Process and Product Quality Assurance.

2.4.7 Software process improvement initiation framework - SPINI

SPINI is an ISO 15504 TR SPICE compatible assessment method (Mäkinen, Varkoi et al. 2000; Varkoi and Makinen 2000). This method has been developed as part of the SataSPIN project which started in August 1998, to establish a software process improvement network (SPIN) in the Satakunta region in Western Finland. The core of the project was to help small and medium sized enterprises SMEs in the software business to develop their operations using international software process models.

The SataSPIN project tailored the SPI initiation phases according to the needs of the participating organisations. The SPI initiation framework consists of three steps:

1. First, the organisation needs to understand the possibilities of SPI in achieving its business goals.
2. Second, the software processes are assessed.
3. Third, the SPI activities need to be planned and supported.

The assessment process has several steps including: holding start-up session, reviewing work product, holding assessment session of two hours on average, reporting results and finally, holding feedback session.

2.4.8 A Modular software process Mini-Assessment method

The modular Mini Assessment MMA method has been developed at Kodak to fulfil the needs of the SPI community in the company to have a common assessment method that uses a standard set of tools and procedures. The objective was to construct a mini-assessment architecture that could be tailored to each project's improvement objectives, life-cycle status, team size and time constraints (Wiegers and Sturzenberger 2000).

The method consists of 8 main steps: planning, opening meeting, CMM orientation, questionnaire administration, questionnaire response analysis, participant discussions, findings generation and findings presentation.

2.4.9 S^{3m} Mini-Assessment method - S^{3mAssess}

S^{3m} mini-assessment method (April, Abran et al. 2004; Paquette, April et al. 2006) has been developed to assess software maintenance processes based on S^{3m} model (April, Hayes et al. 2005; April and Abran 2008). S^{3mAssess} was developed to obtain a reliable maturity rating for maintenance processes without investing too much effort. Additionally, individual assessment components can be selected to focus the investigation on specific concerns and to scope the assessment and rating effort to a level relevant to software maintenance organizations.

A new version of the S^{3mAssess} mini-assessment method has been recently developed based on the research work of (Vincent 2008), and a new assessment tool for S^{3mAssess} has been developed in (TOMASO 2008).

2.5 Summary

The existing lightweight assessment models show some differences and similarities regarding their achievements and processes. A comparison between most of the lightweight assessments in terms of their achievements and processes is available in (Pikkarainen 2006). This comparison is shown in Table 2.2 where, for example, one can see that the MARES

method has three main processes: Planning, Contextualization and Assessment, as addressed previously.

The following points summarize the findings regarding software process assessment.

1. In literature review, most of the publications that discuss problems facing the SMEs, and VSEs when conducting SPI initiatives mention the same points in terms of special organization structure, shortage of recourses, cost and long term ROI associated with SPI initiatives. The severity of these problems becomes apparent as the organization's size becomes smaller.
2. The work in the SPI and SPA field either:
 - a. Discuss the SPI and SPA implemented by large organizations where the comprehensive, heavy-weight assessment methods and improvement approaches are used.
 - b. Discuss the lightweight SPA methods and improvement approaches to fit the needs of SME, small or VSE organizations.
3. Consequently, when discussing the assessment methods, one can recognize two main streams in this research field:
 - a. Comprehensive or heavyweight assessment methods used mainly by large organizations.
 - b. Tailored or lightweight assessment methods used by "non large organizations" including SME and VSE.
4. Researchers of lightweight SPA methods usually alternate between SME and VSE during their discussion of the same SPA method which means that there is no distinct difference between the needs and requirements of assessments used by SME or VSE organizations. Thus, the classification of assessments as comprehensive heavyweight SPA methods and tailored lightweight SPA methods is prominent.

5. The current research trend in software process assessment in SME organizations is to provide different SPA methods without paying attention to objectively evaluate the success of such methods or to what degree these methods fulfill the requirements of engineering design principles.

Based on that premise, the proposed evaluation methods presented in this thesis are used to evaluate lightweight assessment methods which are used by SME or VSE organizations.

Table 2.2

Comparison of the principles and processes of lightweight assessment methods
(Pikkarainen 2006)

Method	Achievement of the evaluation method	Process
MARES	Low assessment costs, detailed description of assessment process and model, conformity with ISO 15504, no specific knowledge required from company representatives, public availability.	Planning, contextualization, process assessment.
RAPID	Low costs, detailed description of assessment model, conformity with ISO 15504, no specific knowledge required from company representatives.	Organization, demographic creation, questionnaire, preparation of assessment plan, preparation of assessment instrument, assessment conduction, assessment report creation.
EPA	Low costs, detailed description of assessment process, conformity with ISO 15504, no specific knowledge required from company representatives.	Select process areas, develop appraisal schedule, conduct over briefing, site briefing, analyze key documents, examine and document objective evidence, generate appraisal results, presentation of the findings report.
SPINI	Detailed description of assessment process and model, conformity with ISO 15504, no specific knowledge required from company representatives.	Needs analyzing, process assessment.

Table 2.2

Comparison of the principles and processes of lightweight assessment methods
(Continued)

Method	Achievement of the evaluation method	Process
FAME	Support for high level process modeling, conformity with ISO 15504.	Not available.
TOPS	Low costs, detailed definition of assessment model, conformity with ISO 15504, public.	Stimulate interests in assessment and improvement, definition of improvement plans, data collection, database creation.

CHAPTER 3

RESEARCH OBJECTIVES AND METHODOLOGY

3.1 Introduction

The SPI process in SME organizations is challenging due several problems facing this type of organization, as discussed in section 2.3. Hence, pursuing an SPI process and discovering that its objectives have not been achieved is a significant waste of limited resources for SME organizations. Pursuing an SPI process and failing is not uncommon. Failures occur in two-thirds of the cases which have pursued a CMM-based SPI (Curtis 1994). Moreover, the cost-effectiveness of process assessment has not been convincingly demonstrated (Gray and Smith 1998) and the rigorous assessments provided by well-known SPI approaches are considered by many small software development firms to be too expensive (Cater-Steel 2004).

Therefore, until new success-guaranteed assessment methods and improvement processes are available, how can the success rate of current methods and of new methods to be increased based on the same concepts used today?

The assessment methods which tailor more comprehensive approaches to fit the needs of SME organizations are generally known as “lightweight assessment methods”. The reliability and effectiveness of comprehensive assessment methods, e.g. CMMI and ISO 15504 compliant SPA methods, have been studied by different researchers, see (El-Emam, Briand et al. 1996; El-Emam and Goldenson 1999; El-Emam and Madhavji 1999). Unfortunately, for lightweight assessment methods used by SME organizations, such studies are not found. The question is how to increase the success rate for lightweight assessments conducted by SMEs in the absence of their reliability studies.

3.2 Research motivation

The authors of lightweight assessment methods typically claim that their assessment methods are successful (Cignoni 1999; Kuvaja, Palo et al. 1999; Rout, Tuffley et al. 2000; Anacleto, Wangenheim et al. 2004; Alexandre, Renault et al. 2006). The authors' claims are based on a few case studies using their proposed assessment methods and also on subjective feedback from organizations that typically indicate their satisfaction with the results obtained as well as providing some comments. Unfortunately, no formal evaluation of these claims has yet been documented. In this thesis the formal development of an evaluation method refers to the use of formal evaluation theory concepts to develop a formal evaluation method. Furthermore, there has been no systematic attempt to synthesize and organize the available data provided by the SPA literature on the evaluation of these methods. Similarly, research in the SPA field to date does not refer to any theoretical justification in the design of the SPA method.

As a result, the motivation for this research project can be summarized with the following three points:

1. There is no independent evaluation of SPA designers' claims that their assessment methods are conducted successfully.
2. There has been no systematic attempt to synthesize and organize the available experiences provided by the literature in the SPA field in terms of conducting assessments.
3. Research in the evaluation of assessment methods to date for SMEs organizations does not refer to any theoretical justification. The theoretical justifications in this context refer to the use of the evaluation theory concepts to build evaluation method to evaluate lightweight SPA methods. (Ares, Garcia et al. 2000) have used the evaluation theory concepts to build an assessment method where the terms "software process assessment" and "software process evaluation" are used interchangeably. Hence, the evaluation

theory concepts are used in (Ares, Garcia et al. 2000) to study the software development process but not the assessment methods development process.

3.3 Research goal

Evaluating the design of SPA methods used by SME organizations and their success opportunities is vital to initiate a successful SPI process. The literature review conducted in this research project has not revealed any comprehensive and formal evaluation method either for evaluating the success of the SPA methods or for studying the design of SPA methods from an engineering design viewpoint. Evaluation methods should be far from being based mostly on an ad hoc approach. The evaluation methods should be developed formally, based on well-founded and empirically verified evaluation concepts. Developing formal evaluation method to evaluate SPA methods would contribute to the enhancement of the process of theory building and to their rigor in the evaluation of software process assessment methods. Accordingly, the goal of this research project can be summarized as:

- **Evaluate the success of lightweight software process assessment methods.**

One should be noted that evaluation is not the same as comparison. Usually, a comparison is conducted at a higher level with fewer details than an evaluation. Although a few comparison methods have been identified in the literature to compare several lightweight SPA methods based on a set of proposed comparison criteria as in (Anacleto, Wangenheim et al. 2004), such comparison methods provide few informational details, such as information on the form: satisfy, do not satisfy, more or less satisfy. Moreover, such comparison methods can not be considered as a formal evaluation of the success of SPA methods since they are not based on evaluation theory concepts. The comparison method used in (Anacleto, Wangenheim et al. 2004) has also been used by (McCaffery, McFall et al. 2005) to compare their own SPA method with other methods. For comparison purposes, the comparison methods found in the literature can be enhanced by making them more informative thereby providing more useful data for those who have conducted the comparison process.

3.4 Research objectives

To achieve the specified goal stated in the pervious section, two objectives for this research project have been selected:

1. To develop a method to evaluate, from an engineering design viewpoint, lightweight software process assessment methods, referred to as the top-down approach.
2. To develop a method to evaluate, based on success evidences found in the literature, lightweight software process assessment methods, referred to as the bottom-up approach.

The research approach selected to pursue these objectives is based on the findings from other disciplines outside software engineering, such as evaluation theory concepts to provide the framework for the proposed evaluation and engineering design classifications including Vincenti's classifications. The research approach also makes use of the experiences of researchers and practitioners in the field of SPA methods to collect the success evidences that affect lightweight SPA methods. The referenced disciplines are shown in Figure 3.1, and will be presented in further detail in chapter 5.

Finally, the work conducted in this research project is different from the similar work of (Komi-Sevio 2004), where an evaluation framework for the SPI methods is proposed. This work specializes in the study of SPA methods separately from the improvement process. This separation has made it possible to study the assessment methods in detail while other studies mentioned the process assessment implicitly as part of the SPI initiative and hence is not studied thoroughly.

3.5 Research scope

This research project is concerned with evaluating SPA methods which are part of top-down SPI approaches including model-based approaches. Although the evaluation methods

developed in this research project can be applied to any SPA method, the focus of this research is on evaluating lightweight SPA methods, since there is a lack of studies of the effectiveness and performance of lightweight SPA methods.

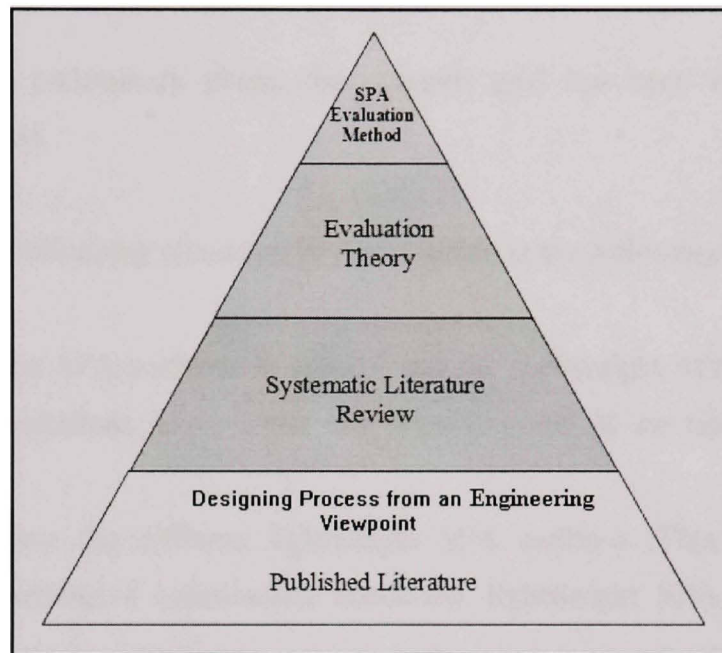


Figure 3.1 Summary of the reference disciplines.

3.6 Research methodology

The research methodology is divided into three phases, each phase consisting of several steps. The three phases are:

1. Preliminary phase

The preliminary phase focuses on exploring the research directions related to the software process assessment and improvement in general and on SPA in SME and VSE organizations in particular. In this phase, the different SPI frameworks and the different assessment methods are also explored. The problems facing SME and VSE organizations and the lightweight SPA methods used by these organizations are also explored. The key findings of

this phase were presented in chapter 1 and chapter 2. These findings are used to formalize the research goal and objectives addressed in chapter 3. The preliminary phase is finalized by summarizing the key points for each lightweight SPA method and comparing them; a comparison framework is proposed and presented in chapter 4.

At the end of the preliminary phase, the research goal has been formulated, research objectives established.

Consequently, the preliminary phase can be summarized in the following steps:

- Step1: Study the SPA methods in general and the lightweight SPA methods for SME and VSE organizations in particular and identify some of the research issues in this field.
- Step 2: Compare the different lightweight SPA methods. This comparison would provide comprehensive information about the lightweight SPA methods currently available.
- Step 3: Identify research goal and select the research objectives.

2. Development phase

The development phase works on achieving the specified research objectives in developing evaluation methods to evaluate, based on an engineering viewpoint, the design of the lightweight SPA methods and to evaluate, based on systematic literature review, the success of the lightweight SPA methods.

The development phase can be divided into the following steps, in a continuous sequence with the previous steps in phase 1:

- Step 4: Identify the reference disciplines.

The two main disciplines referenced in this phase are:

- The evaluation theory which will be used as a framework to develop both the bottom-up and top-down evaluation methods;
- Engineering design classifications which will be used to develop the top-down evaluation method.
- Step 5: Develop the Top-down evaluation method.
In this evaluation method, the engineering design classifications proposed by (Vincenti 1990) will be used as bases to identify the criteria that should be considered during the design of lightweight SPA methods. Vincenti's classifications are used to verify to what extent the design of these methods align with the engineering design principles.
- Step 6: Develop the Bottom-up evaluation method.
In this evaluation method, a systematic literature review will be conducted based on the guidelines in (Kitchenham 2007). The literature review aims at collecting a set of success evidences published as lessons learned, success factors, requirements and observations. The collected set of success evidences will be the core for the evaluation method.
- Step 7: Verify the development process of the two evaluation methods.
This step will verify the research process by using a set of verification strategies to contribute to the validity of this research project. "Strategies of verification are those techniques that contribute to the validity of the project and are implemented in the actual research process" (Morse, Swanson et al. 2001).

3. Testing phase

Testing the developed evaluation methods is step 8 in this research methodology. The developed evaluation methods will be tested by conducting three case studies to evaluate three different lightweight SPA methods. The three SPA methods were developed in three different countries. The three methods to be tested will be the OWPL Micro-Evaluation method developed in Belgium (Alexandre, Renault et al. 2006), the MARES method developed in Brazil (Wangenheim, Anacleto et al. 2004) and the S3^{mAssess} method developed in Canada and used to assess the software maintenance process (April and Abran 2008).

Finally, the resulting evaluation methods will be presented and released in step 9, to be used by those interested in evaluating lightweight SPA methods including practitioners from the industry and the designers of new lightweight SPA methods.

Figure 3.2 summarizes the three phases and the 9 steps of the proposed research methodology.

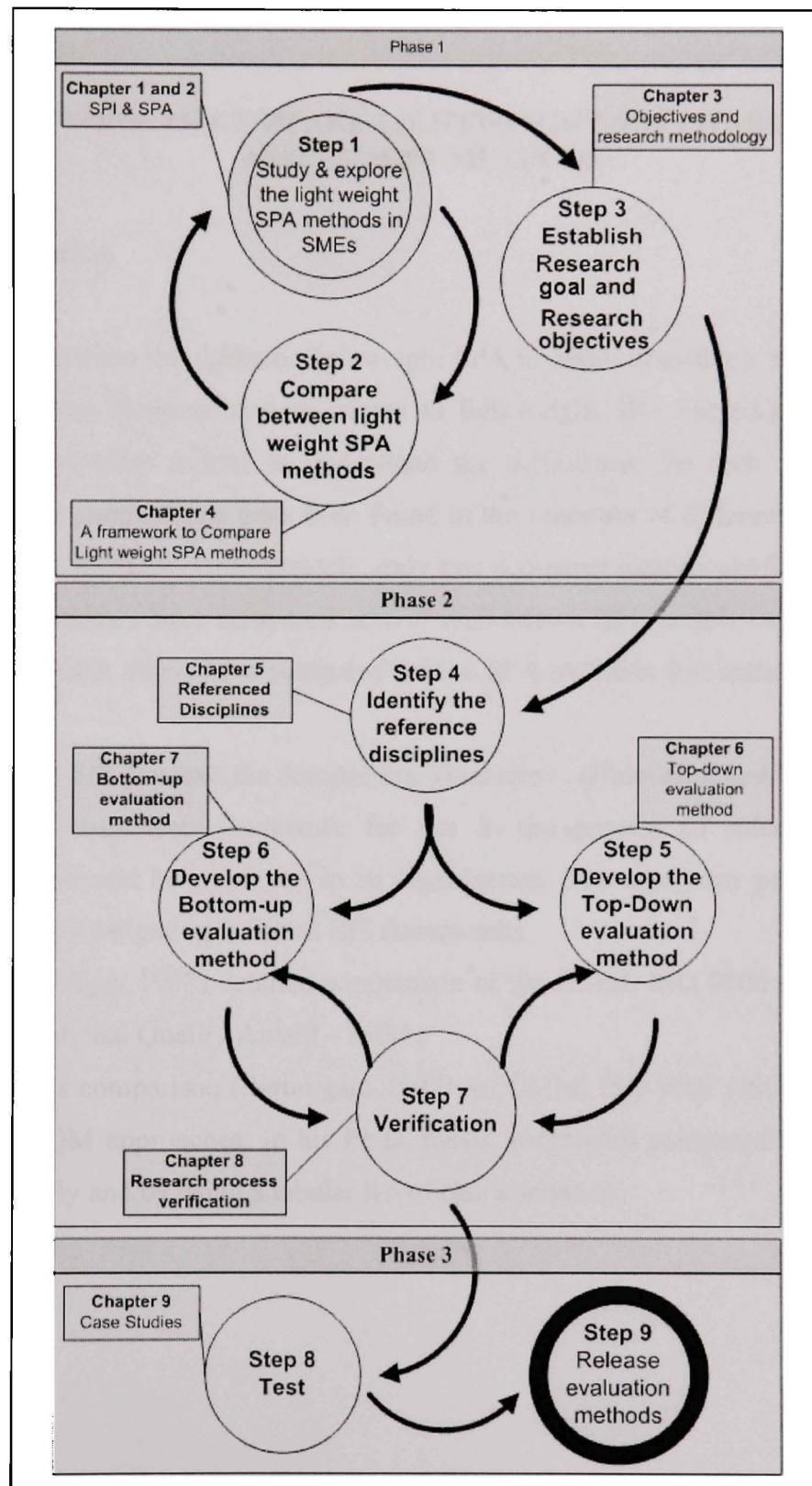


Figure 3.2 Research methodology.

CHAPTER 4

A FRAMEWORK TO COMPARE LIGHTWEIGHT SOFTWARE PROCESS ASSESSMENT METHODS

4.1 Introduction

A comparison between the different lightweight SPA methods provides a tool to summarize the findings of the literature review related to lightweight SPA methods as discussed in chapter 2 and provides a tool to understand the differences between these assessment methods. Several comparisons have been found in the literature of different SPI models and their related SPA methods. Unfortunately, only one compares lightweight SPA methods. For example, some authors have compared several well-known SPI models such as CMM, ISO 15504 and ISO 9000; others have compared several SPA methods. For instance:

1. Halvorsen's SPI frameworks comparison Taxonomy (Halvorsen and Reidar 1999): A high-level comparison framework for use in the process of selecting which SPI framework should be employed in an organization. The taxonomy points out areas of interest for investigating different SPI frameworks.
2. Tingey's (Tingey 1997) detailed comparison of the CMM, ISO 9000 and the Malcolm Baldrige National Quality Award – MBA.
3. Sørungård's comparison (Sørungård 1997) of CMM, ISO 9001 (ISO-9001 1993) and QIP/EIF/GQM approaches. In his Ph.D. thesis Sørungård compared these approaches both textually and by using a tabular list of characteristics.

4. (El-Emam, Drouin et al. 1998) textual comparison of ISO TR 15504¹ and ISO 9001 (ISO-9001 1993) to show their differences and provides a table showing SPICE-ISO 9001 mapping.
5. Paulk's (Paulk 1995) comparison of ISO 9001 (ISO-9001 1993) and the CMM.
6. Anacleto's (Anacleto, Wangenheim et al. 2004) comparison of several lightweight process assessment methods for small companies.
7. McCaffrey's (McCaffery, McFall et al. 2005) comparison of his proposed assessment method dedicated to SME organizations to other lightweight assessment methods.
8. Pikkarainen's comparison (Pikkarainen 2006) of the principles and processes of lightweight assessment methods used by SME organizations. This comparison was shown in chapter 2, Table 2.2.

In the first five references, the work is dedicated to the comparison of several SPI models either by textual or characteristics comparison among the different models. In the last three references, the authors are more interested in comparing the assessment methods dedicated to SME organizations. McCaffrey in his comparison (McCaffery, McFall et al. 2005) used the same characteristics comparison method proposed by Anacleto (Anacleto, Wangenheim et al. 2004).

Next, two comparison methods are presented in detail: the first is Halvorsen taxonomy to compare SPI frameworks (Halvorsen and Reidar 1999), and the second is Anacleto et al.'s method (Anacleto, Wangenheim et al. 2004).

¹ ISO 15504 (1998), which is also known as SPICE, has been published as a document of 9 parts and has been revised extensively and republished from 2003 to 2007 as an international standard of 5 parts.

4.2 Comparison methods in detail

This section presents two comparison methods in detail, the first one compares different SPI frameworks and the second one compares lightweight SPA methods.

4.2.1 Halvorsen's taxonomy

Halvorsen recognized four different classes of methods for comparing SPI frameworks; “From our review of other comparison work we have recognized four main classes of comparison methods” (Halvorsen and Reidar 1999), which are:

1. Characteristics comparison method:

The comparison in this method is based on a set of predefined characteristics listed in tabular form and gives a compact and high-level comparison method with few details.

2. Framework mapping comparison method:

This method is the process of creating a map from the statements or concepts of one framework to those of another. This approach is useful when an organization employs two or more different SPI frameworks, as corresponding statements can be identified and redundancy reduced. Thus, the extra effort needed to employ more than one framework is minimized.

3. Bilateral comparison method:

In a bilateral comparison, two frameworks are compared textually. The difference between this comparison method and the two previously noted is its textual nature.

4. Needs mapping comparison method:

The Needs mapping method does not constitute a direct comparison of frameworks, but considers the organizational and environmental needs that must be considered when selecting which SPI framework to adopt.

Halvorsen has proposed a taxonomy which falls into the “Characteristics comparison method” grouping described above. In Halvorsen's taxonomy, 25 different characteristics are

defined to compare SPI frameworks. Within the taxonomy, the characteristics are grouped into 5 categories, as shown in Table 4.1.

Halvorsen used his taxonomy to compare six SPI frameworks: TQM, CMM v1.1, ISO 9000, ISO 15504 1998, GQM and SPIQ. For more details on the characteristics and comparison results, refer to (Halvorsen and Reidar 1999).

Table 4.1

Categorization of characteristics in Halvorsen taxonomy
(Halvorsen and Reidar 1999)

General	Process	Organization	Quality	Result
<ul style="list-style-type: none"> • Geographic origin/ spread • Scientific origin • Development/ stability • Popularity • Software specific • Prescriptive/ descriptive • Adaptive 	<ul style="list-style-type: none"> • Assessment • Assessor • Proc. Improve. method • Improve. initiation focus • Analysis techniques 	<ul style="list-style-type: none"> • Actors/roles stakeholders • Organization size • Coherence 	<ul style="list-style-type: none"> • Quality perspective • Progression • Causal relation • Comparative 	<ul style="list-style-type: none"> • Goal • Process artifacts • Certification • Cost implement • Validation

In general and for the purposes of comparing different lightweight SPA methods which are built based on an SPI framework, Halvorsen's taxonomy can still be used with some adaptation. Indeed, one observes that several characteristics can be discarded from Halvorsen's taxonomy since their values are implicitly known or inherited directly from the underlying assessment framework. The candidate characteristics to be discarded will be discussed later.

4.2.2 Anacleto comparison method

Anacleto et al. (Anacleto, Wangenheim et al. 2004) have proposed a tabular comparison of five different assessment methods for small organizations namely: RAPID, SPINI, FAME

and TOPS, and their new method, MARES. Their comparison is based on the following criteria:

1. Low cost.
2. Reliable results.
3. Detailed description of the assessment process.
4. Guidance for process selection.
5. Detailed definition of the assessment method.
6. Support for identification of risks and suggestions for improvement.
7. Support for high-level process modeling.
8. Conformity with ISO 15504 (version not specified in the source).
9. No specific software engineering knowledge required from the company representative;
10. Tool support.
11. Integrated into the assessment methodology.
12. Public availability.

Based on the Halvorsen's classification, this comparison method also falls into the "Characteristics comparison method" group. Anacleto et al. did not aim to define a comparison framework; rather, they documented their own criteria for comparing several assessment methods to verify to what extent their proposed MARES assessment method is similar to other methods. On the basis of these criteria, Anacleto et al. concluded that their method is most similar to the SPINI method. Moreover, the values assigned to the above characteristics provide few details. The values are: satisfy; do not satisfy; more or less satisfy and no information available. Although these criteria are relevant for comparison purposes, other criteria can be added to compare different lightweight assessment methods including RAPID, SPINI, FAME, TOPS, Micro-Evaluation and MARES to convey more informative and useful data.

As mentioned previously, some characteristics of Halvorsen's taxonomy can be discarded. Some explanations for discarding such characteristics are presented in Table 4.2. The values

of these unnecessary characteristics did not add new knowledge for those performing the comparison and increase the time of the comparison process.

Table 4.2

Descriptions of some examples for excluded characteristics

Discarded Characteristic	Explanation
Software specific	All the SPI frameworks and methods discussed for SME organization are software specific so this attribute is implicitly known to be 'yes' for all; and is therefore excluded.
Prescriptive/ Descriptive	The SPI frameworks are normally both prescriptive and descriptive and hence no need to check this characteristic.
Adaptability	Other characteristics such as the number of processes assessed, the name of these processes and the guidance for process selection to be improved give an indication of the adaptability of the assessment method and to what degree it fulfills customer needs. No need to add extra characteristics to check adaptability.
Assessment	The software process assessment is usually conducted to determine the process capability and/or organization maturity. Since this characteristic is known implicitly there is no need to be mentioned in a separate characteristic.
Assessor	The software process assessment assessor can be either internal or external. Internal assessors are chosen for conducting self-assessment and external assessors are chosen for certification or process improvement purposes. Therefore this characteristic has no significant value as a comparison characteristic.

As a result, a comparison framework dedicated for SME organizations could be designed by merging Halvorsen's Taxonomy and Anacleto's comparison methods and proposing some other useful characteristics for SPA methods for SME organizations, see Table 4.3.

4.3 Proposed comparison framework

A comparison framework that is specifically useful for comparing lightweight SPA methods is proposed in this section. According to the Halvorsen's classification of comparison

methods, the proposed method will belong to the “Characteristics comparison method” grouping.

This grouping includes several characteristics from both the Halvorsen and Anacleto et al. comparison frameworks to make them more convenient for use by SME organizations. Several characteristics that do not have strong informative value or that are common to SPA methods, have been excluded, as discussed in the previous section, since the presentation of the characteristics that show the differences between the compared methods is more vital and can help in deciding which method to use. Other criteria involving useful and informative data about the methods compared have been added, including how long the assessment should take, the number of assessed processes and what they are, as well as which of these processes are to be improved upon. The set of characteristics to be used in the proposed comparison framework is shown in Table 4.3

Table 4.3

Categorization of characteristics in the proposed comparison framework

Halvorsen Characteristics	Anacleto et al. Characteristics	New Characteristics
<ul style="list-style-type: none"> • Geographic origin/spread • Scientific origin • Development/stability • Popularity • Analysis techniques 	<ul style="list-style-type: none"> • Cost • Guidance for process selection • Support for identification of risk and improvement suggestions • Need for specific SE knowledge from the company representative • Tool support • Public availability • Detailed description of assessment process • Detailed definition of assessment model 	<ul style="list-style-type: none"> • Number of assessed processes • Assessed processes • Number of processes to be improved • Assessment duration

4.3.1 Characteristics description

1. Geographic Origin/Spread: Where did the framework originate and where is it used now?
2. Scientific Origin: What is the background upon which this framework is based?
3. Development/Stability: one should employ an evolved and relatively stable framework which is achieved through experience and real use over a number of years?
4. Popularity: Where is this method used?
5. Analysis techniques: Does the framework utilize any quantitative or qualitative analysis techniques, such as statistical process control (SPC) or questionnaires?
6. Cost: What is the relative cost? Is this cost high, low or moderate?
7. Guidance for process selection: Does the compared method provide any guidance in selecting the processes to be assessed?
8. Support for identification of risk and suggestions for improvement: Does the compared method provide a way to identify the possible risks and suggestions for improvement?
9. Need for specific SE knowledge from the company representative: Does the company representative need to have any particular level of software engineering knowledge?
10. Tool support: What tools does the assessment method use to support the assessment process?
11. Public availability: Is the compared method available to the public?
12. Detailed description of assessment process: Does the assessment method provide a detailed description of the assessment process?
13. Detailed Definition of assessment model: Does the assessment method provide a definition of the underlying assessment model?
14. Number of assessed process: What is the number of processes to be assessed by this framework?
15. Assessed process: What processes have been chosen for assessment?
16. Number of processes to be improved: From the set of assessed processes, how many will need to undergo improvement?
17. Assessment duration: How long will the assessment process take?

4.3.2 Application of the comparison framework

As mentioned previously, the comparison between different lightweight SPA methods is useful to summarize and understand these methods. This comparison helps organizations and practitioners interested in initiating an assessment process to choose the SPA method that fits their needs and goals. The comparison result of seven different lightweight SPA methods is presented in Table 4.4 at the end of this section. The compared methods are:

1. TOPS (Cignoni 1999): Toward Organized Process in SMEs.
2. Micro-Evaluation (Habra, Renault et al. 2002): OWPL Micro assessment method.
3. MARES (Wangenheim, Anacleto et al. 2004): A methodology for software process assessment in small software companies.
4. SPM (Richardson 2001): Software Process Matrix.
5. RAPID (Rout, Tuffley et al. 2000): Rapid Assessment for Process Improvement for Software Development.
6. FAME (Beitz, Emam et al. 1999): Fraunhofer Assessment Method.
7. EAP (McCaffery, McFall et al. 2005): Improving the Express Process Appraisal Method.

When there is no information that can be used to answer any of the specified criteria, the value for that criterion is stated as not available (NA).

Table 4.4

The comparison results of the seven lightweight SPA methods

Criteria	MARES	TOPS	FAME	RAPID	SPM	EAP	Micro-Evaluation
Geographic origin/Spread	Brazil	Italy	Germany	Australia	Ireland	Ireland	Belgium
Scientific origin	ISO 15504	ISO 15504	ISO 15504/ Bootstrap	ISO 15504	Quality Function Deployment	CMMI Compliant with the ARC 1.1	OWPL
Cost	Low	Low	NA	Low	Low	Low	Low
Development/ Stability	NA	NA	NA	Since 1999	Since 1999	Since 2003	Since 1998
Application region	Regional	Regional	Regional	Regional	Regional	NA	Belgium/ Quebec/ France
Analysis techniques	Interview	Interview	Interview	Interview	Questionnaire	Interview	Short Interview
Provide detailed description of assessment process	Totally	Partially	NA	Partially	NA	Yes	Partially
Provide detailed definition of assessment model	Yes	No	Yes	No	NA	Partially/ Make reference to CMMI	Yes
Number of processes assessed	26	3	4	8	47 Process with 135 practices	6	6
Number of processes to be improved	2-3	3	4	8	Max. 10 practices	6	6
Assessment duration	1 day	Half a day	NA	1 day	NA	1 day	Half an hour
Guidance for process selection	Yes, uses a contextual. Phase	No	NA	No	NA	No	No

Table 4.4

The comparison results of the seven lightweight SPA methods (Continued)

Criteria	MARES	TOPS	FAME	RAPID	SPM	EAP	Micro-Evaluation
Assessed processes	Selected after identifying strengths and weakness based on SWOT analysis	1. Req. Analysis 2. Ver. & test 3. Joint review	NA	1. Req. gathering 2. Software development 3. Project management 4. Configuration management 5. Quality assurance 6. Problem resolution 7. Risk management 8. Process establishment	Selected according to a prioritized list based on QFD calculations	1-Requirement Management 2-Configuration Management 3-Project Planning 4-Project Management 5-Project monitoring and control 6-Process & Product QA	1. Quality Assurance 2. Customer Management 3. Supplier Management 4. Project Management 5. Product Management 6. Training and Human Resource Management
Tool support	NA	Paper forms	Data collection, analyses and rating tools	Paper Forms	NA	Paper forms + data collection & analysis tools	Paper forms + Excel sheet
Public availability	Yes	Yes	No	No	NA	No	No
Support for identification of risk and improve. Suggestions	Yes, uses a risk manage. Phase	Partially supported	Partially supported	No	NA	Yes	Partially supported
Need for specific SE knowledge on the part of the company representative	No	Yes	Yes	Yes	NA	No	No

CHAPTER 5

REFERENCE DISCIPLINES

This chapter presents the reference disciplines which are used as a basis to develop the proposed evaluation methods; firstly, the evaluation framework used to build the evaluation methods is presented and secondly, different views of technology in general and software technology in particular are presented.

5.1 Evaluation theory concepts

An evaluation can be defined as the process of determining merit, worth, or significance (Scriven 1991). However, there are also common synonyms for the terms in this definition: “quality” is often used instead of “merit,” “value” instead of “worth,” and “importance” instead of “significance.” Also, due to the fact that the evaluation is an anxiety-provoking activity for most people, other synonyms for the term can be found: analysis appraisal, audit, review, examination (Scriven 2003).

(Scriven 2003) provided a transdisciplinary vision - such as logic, design, and statistics for the future of evaluation science:

“I hope and expect that the essential nature of evaluation itself will crystallize in our minds into a clear and essentially universal recognition of it as a discipline, a discipline with a clear definition, subject matter, logical structure, and multiple fields of application. In particular, it will, I think, become recognized as one of the elite group of disciplines, which I call transdisciplines. These disciplines are notable because they supply essential tools for other disciplines, while retaining an autonomous structure and research effort of their own” (Scriven 2003).

Although there are diverse theories of program evaluation (Shadish, Cook et al. 1993) which resulted in not developing a general theory of evaluation common to all disciplines, this diversity helps evaluators identify the components needed to execute an evaluation.

In the software engineering (SE) field, usually, evaluations are developed and performed without taking into account the efforts and lessons learned in other software and non-software disciplines (López 2003). However, the study of evaluation theories and methods already developed in other disciplines such as in (Scriven 1991; Shadish, Cook et al. 1993; Worthen, Sanders et al. 1997) could help to elaborate more detailed, complete and systematic evaluation methods for application in the diverse SE areas (López 2003). An analysis of different evaluation approaches and generalization of the evaluation elements were carried out by a few researchers in the SE field, e.g. (ACUÑA, ANTONIO et al. 2000; Ares, García et al. 2000; López 2000; López 2003), with the aim of identifying a set of basic elements common to any type of evaluation method. The classification of the evaluation methods is discussed in section 5.1.1 and the basic components for evaluation are discussed in section 5.1.2.

5.1.1 Evaluation method classifications

As a result of the diversity in theories of program evaluation, several classifications of evaluation methods exist; for instance, (Worthen, Sanders et al. 1997) classified the evaluation methods into six approaches shown in Table 5.1.

Another classification proposed by House (House 1980) divided the evaluation methods, based on the type of results that will be obtained, into objective and subjective methods as shown in Table 5.2

Table 5.1

Evaluation methods proposed by Worthen
(Worthen, Sanders et al. 1997)

Evaluation Approach	General Purpose of Evaluation
Objective-oriented evaluation	Determining the extent to which goals are achieved
Management-oriented evaluation	Providing useful information to aid in making decisions

Table 5.1

Evaluation methods proposed by Worthen (Continued)

Evaluation Approach	General Purpose of Evaluation
Consumer-oriented evaluation	Providing information about products to aid in making decisions about purchases or adoptions
Expertise-oriented evaluation	Providing professional judgment of quality
Adversary-oriented evaluation	Providing a balanced examination of all sides of controversial issues, highlighting both strengths and weaknesses
Participant-oriented evaluation	Understanding and portraying the complexities of a programmatic activity, responding to an audience's requirements for information

The diverse components described for each type of evaluation has been analyzed by (López 2000) and a set of elements is obtained that can be classified as basic because they are common to any type of evaluation method. These basic components are the topic of the following section.

Table 5.2

Evaluation methods proposed by House
(House 1980)

Objective Methods	Objective-oriented evaluation
	Decision-making evaluation
	Need-oriented evaluation
	Efficiency-oriented evaluation
	Control-oriented evaluation
Subjective Methods	Method of negotiation
	Criticism of art
	Dialectic method

For each type of evaluation methods considered, all the evaluation components should be defined explicitly for the evaluation to be conducted rigorously (López 2000).

5.1.2 Basic components of an evaluation

The six basic components of an evaluation are shown in Figure 5.1. When designing an evaluation method, the six evaluation components should be developed.

The evaluation components, can be defined as (Scriven 1991):

1. Target: The object under evaluation.
2. Criteria: The characteristics of the target which are to be evaluated.
3. Yardstick: The ideal target against which the real target is to be compared.
4. Assessment techniques/Data gathering techniques: The techniques needed to assess each criterion under analysis.
5. Synthesis techniques: Techniques used to organize and synthesize the information obtained with the assessment techniques. The result of the synthesis is compared with the yardstick.
6. Evaluation process: A series of activities and tasks by means of which an evaluation is performed.

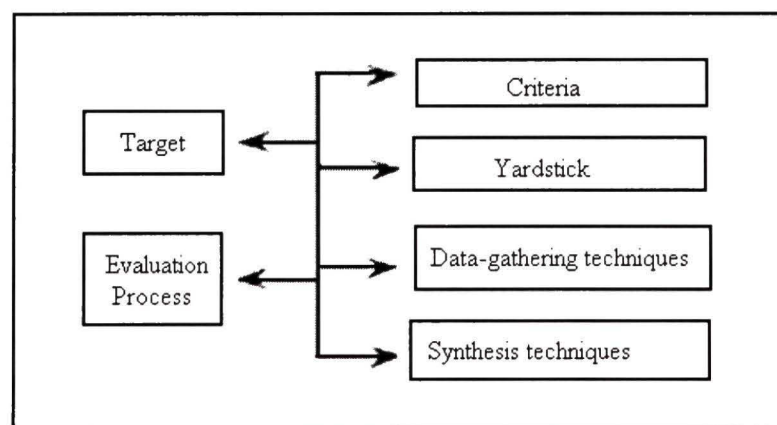


Figure 5.1 Components of an evaluation.
(Scriven 1991)

As shown in Figure 5.2, all these components are closely interrelated. The evaluation can be customized by means of the target, because this is one of the parameters used to select the

evaluation method. Once the target is known and delimited, its characteristics must be identified for evaluation (ACUÑA, ANTONIO et al. 2000; López 2000). All the characteristics and their ideal values, which indicate what the target should be like under ideal conditions, make up what is known as the yardstick or standard. Data about the real target should be obtained using certain data-gathering techniques: a value (numerical, data, information set, etc.) will be gathered for and assigned to each criterion. The data, once collected, are organized in an appropriate structure and compared against the yardstick by applying synthesis techniques. This comparison will output the results of the evaluation. Finally, the entire components above are linked together through the evaluation process (López 2000).

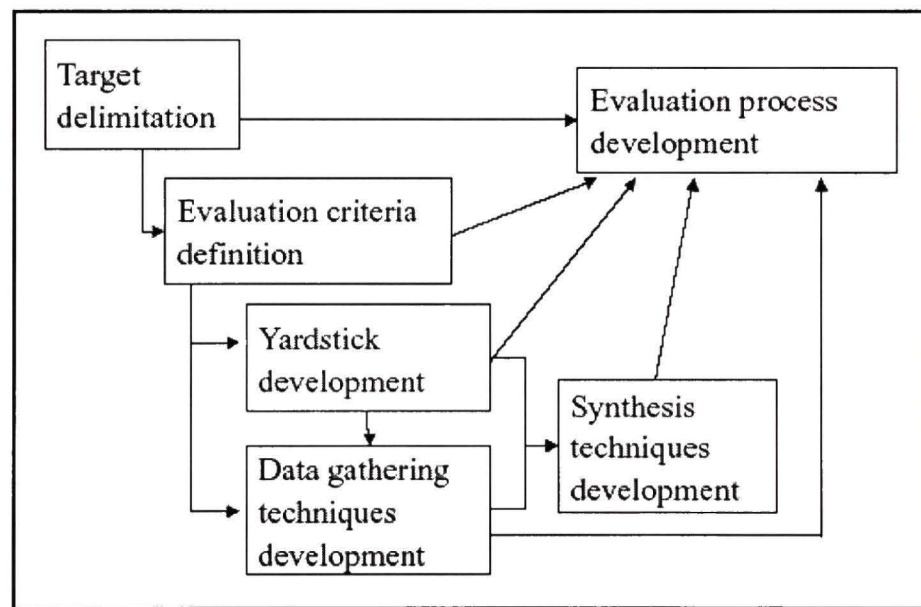


Figure 5.2 Interrelations between components.
(López 2000)

5.2 Overview of the evaluation components in the SPA context

When applying the evaluation theory to any context, one should decide which evaluation method to use taking into consideration the target to be evaluated. In the SPA method evaluation context and based on House classification (House 1980) shown in Table 5.2, the objective method named “control-oriented method” has been selected, since the objective of

this method is to ensure that the target is controlled and directed by the yardstick specified. In the SPA context, the successful design and implementation of SPA method should be controlled and directed by the yardstick developed in this research.

The six evaluation components, shown in Figure 5.1, form the basis for the evaluation framework. An evaluation method can be developed by instantiating the evaluation framework. If the framework is used by other evaluators the resulted evaluation method may not necessarily be exactly the same, since the instantiation process involves making a series of decisions which may differ depending on the opinions and the environment of the evaluation method developer (López 2000).

5.2.1 Target

The first step in defining an evaluation framework is delimiting the evaluation target. To be able to identify the criteria evaluation component, it is necessary to study and delimit the object under evaluation which means identifying the factors to be considered (López 2000). Few guidelines and techniques are available to define the target. “Indeed, it is the experts in the field who very often indicate which factors are to be considered; however, evaluators can apply the functional analysis technique, described as the general description of the target’s function. Depending on what the target is, evaluators can complete this analysis with the description of the context, stage of the development, expected effects, and any other information that can help the evaluator understand what the target is and delimit explicitly what will be analyzed in this evaluation. Sometimes the functional analysis technique is also used to identify the criteria of the evaluation” (López 2000).

In this research project, the SPA method is the target. Building a new SPA method is not the goal of this research project. This research aims at evaluating the design and implementation success of SPA methods, hence, the current SPA methods are studied, giving an evaluation method that can be used to analyze the SPA methods showing their strengths and weaknesses

allowing for improvement which increases the probability of having a successful software process assessment.

5.2.2 Criteria

Defining the criteria is the second essential and critical step in the development process of an evaluation method. In this step, the characteristics of interest for the target are to be defined. These characteristics are referred to as evaluation criteria. Criteria elicitation can be made by either using an obligatory standard that contains implicitly the criteria to be applied in the evaluation or by using other diverse techniques for criteria elicitation. These techniques include (López 2000):

1. Functional analysis of the target defined as the detailed description of the target's function.
2. Needs assessment, referring to any study of the needs, wants, market preferences, values, standards, or ideals that might be relevant to the target.
3. Complex logical analysis when the definition needs more unpacking in order to figure out its implications. This is more often the case when the criterion is significance related. The analysis is a complex inferential process starting from data and definitions.

Figure 5.3 presents a breakdown graph summarizing the different techniques for criteria elicitation. All these techniques can be complemented using the basic set of questions: what, why, when, how, where and who. The purpose is to complete the analysis of the target and to assure that the target has been studied in detail (López 2000). Usually, two types of responses can be gained from these questions:

1. Specific criteria where the characteristics can be assigned a value directly by using certain data-gathering technique.

2. General criteria where characteristics cannot be assigned a value directly and require further decomposition to which the set of questions will be applied successively until specific criteria are obtained.

The collected set of specific and generic criteria can be represented using a diagrammatic tree which is called a criteria tree. This criteria tree forms the basis for developing the evaluation yardstick.

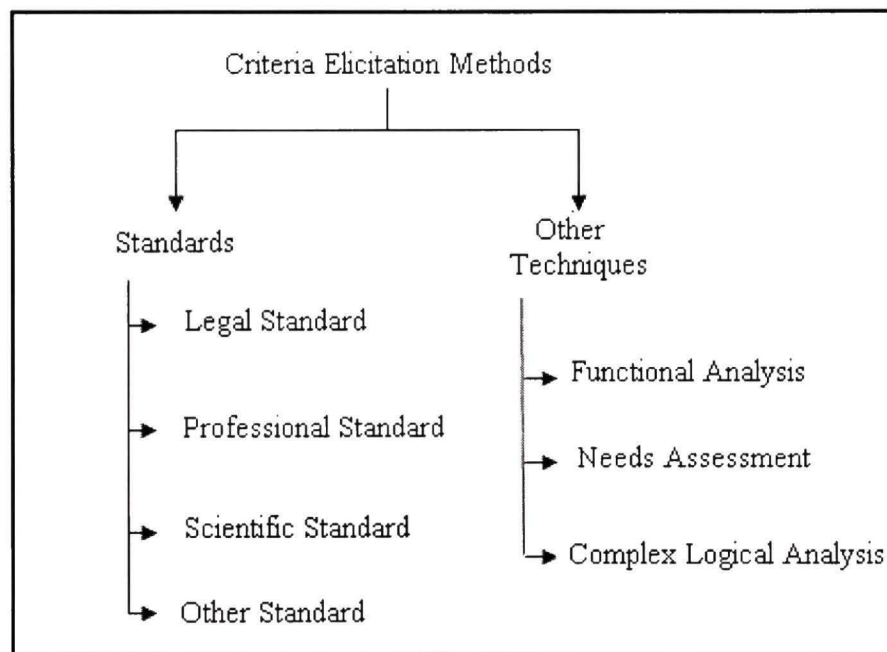


Figure 5.3 Breakdown graph for criteria elicitation methods.

In the SPA process context, the criteria will be developed using two different approaches: the first approach is based on the needs assessments elicitation method. In this approach the SPA method design needs are analyzed based on engineering design principles. This approach and the resulted evaluation method are discussed in chapter 6.

The second approach is also based on the needs assessments elicitation method. In this approach the needs for successful implementation of the SPA method are analyzed. These needs are represented by a set of requirements, success factors and lessons learned which are

gained as a result of implementing the SPA method by different researchers and practitioners in the field. This approach and the resulting evaluation method are discussed in chapter 7.

5.2.3 Yardstick

The target definition and the criteria tree developed in the previous two steps are the basis for developing the yardstick. All yardsticks must contain the specifications, requirements, descriptions, or values for each criterion considered. So, if evaluators have to develop the yardstick, the following bases (López 2000) can be used:

1. The yardstick used in the evaluation should be developed from the criteria tree obtained in the preceding step. The general structure of the yardstick should be inferred from the criteria tree.
2. The yardstick must contain the specifications of all defined criteria.
3. For each criterion, whenever possible, the yardstick must define the specifications structured as pairs [criterion, datum/information].

Whenever applicable, the yardstick must contain threshold values to indicate the minimum value for each criterion to be reached for a positive evaluation. This task is closely related to synthesis techniques. For this reason, these threshold values could be defined when the evaluator has selected the synthesis techniques to be applied in the evaluation (López 2000).

In this research context, the yardstick is based on the evaluation criteria that will be developed in the next two chapters. The yardstick in this case will be very simple, consisting of the different criteria that should be taken into consideration during the design and implementation phases of the SPA method. Hence, the synthesis technique would verify criterion-by-criterion to ensure that each criterion has been considered in the design phase. The threshold value in this case would be fully adequate, partially adequate or not adequate for each criterion. The criterion which is evaluated as partially or non adequate is considered to be a weakness point in the SPA method that may cause its failure.

5.2.4 Data gathering techniques

In this step the applicable data-gathering techniques should be identified. These techniques are used to obtain the information needed to judge the target with the next component. The main data-gathering techniques used in most evaluations in the software engineering field can be classed in three groups (López 2000):

1. Measurement: Involves the use of the appropriate measurement instruments or mechanisms.
2. Assignment: Such as questionnaires, interviews (individual or groups), documentation inspection and simple tests not involving measurement applications.
3. Opinion techniques: Used to get subjective criteria data, such as, by observation.

These techniques are also used in other disciplines. For example, to evaluate social programs, the most common techniques are: survey, interview, test, observation, group techniques, case study, photograph, videotape, slides, document review and analysis, portfolio review, testimonials, expert or peer review, simulated problem or situation, journal or log or diary, and unobtrusive measures (Taylor-Powell and Steel 1996).

The selection of data-gathering techniques will depend on the preceding components (target, criteria, and yardstick). One or more data-gathering techniques must be assigned to each criterion by analyzing the meaning of the criterion and the type of value (numerical, data, etc.) specified for it in the yardstick. Once identified and assigned, each technique must be developed, outputting questionnaires, standard interviews, lists of documents for inspection, observation forms, metrics, to list a few. The attachment of examples of the practical application for each technique (López 2000) is also recommended.

In this research context, the data-gathering techniques selected are assignment techniques. The main techniques used are documentation review and questionnaires. The documents reviewed are the set of published articles and reports that discuss the design and

implementation process as well as the results obtained from the evaluated methods. The questionnaires are filled out by the designers of the evaluated SPA methods. The criteria tree and the yardstick will be used to develop the questionnaire. This questionnaire can be used by the practitioners interested in evaluating SPA methods or by the designers of new SPA methods to ensure all requirements mentioned in the evaluation criteria have been taken into consideration.

5.2.5 Synthesis techniques

Synthesis techniques are used to synthesize all the data and information obtained after applying the data-gathering techniques and for comparison against the yardstick. Usually, two types of synthesis techniques can be applied (López 2000):

1. Single value: A single datum (numerical or otherwise) is obtained as a result of the evaluation. This group includes combination methods. When these techniques are applied, a meaningful value scale is required for the datum obtained.
2. Multiple values: These techniques, for example, statistical techniques, criteria grouping, and datum-by-datum comparison with the yardstick, output more detailed information than single-value techniques.

The selection of the synthesis techniques depends on the previous components. In this research context the multiple values technique is used where the datum-by-datum comparison with the yardstick is applied.

5.2.6 Evaluation process

The evaluation process is a series of specific activities and tasks that are to be executed to perform an evaluation. All the previous components are necessary to describe and design an evaluation method, but it is the evaluation process that describes the list of activities to perform and when to use the previous elements in practice. The framework describes three

main phases, as shown in Figure 5.4. These phases match the three major points through which an evaluation passes:

1. **Planning or preparation:** Activities involving making contact with the target to be evaluated, delimiting the evaluation, and planning and managing its execution. This phase ends when all the evaluation components have been developed (or, if necessary, adapted to the target in question) and the team of evaluators is ready to make a visit to or interact with the target.
2. **Examination:** Application of the data-gathering techniques and obtaining the data required to judge the target. This phase ends when all the information has been obtained for all the criteria considered in the evaluation.
3. **Decision making:** Application of the synthesis techniques and development of the final report. Also, this activity includes the task of completing the documentation of the evaluation, whose goal is double: to refine the evaluation process (and therefore the method of evaluation) and to maintain the documentation of this evaluation to be compared with future evaluations of the same or similar target.

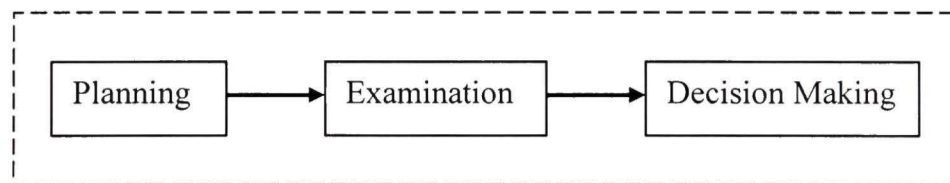


Figure 5.4 Main phases of the evaluation process.

In the SPA context, the activities associated with each phase are shown in Figure 5.5. In the planning phase, the target should be analyzed first. This analysis is needed to be able to take further decisions regarding the different evaluation components such as criteria and yardstick. In the next step, the evaluation components have to be defined and adapted to the SPA process context. The definition of the evaluation components creates the evaluation framework which is instantiated to create the evaluation method. The instantiation includes defining the criteria, yardstick, assessment and synthesis techniques. At this stage the SPA

methods documentation and any other information that would be useful for conducting the evaluation should be prepared.

In the examination phase, the evaluator starts by assuring that all necessary data for the evaluation are available and complete. Then, the data gathering technique is applied to collect the data and verify that the data collected is complete and answers all the questions in the questionnaire.

The final stage starts by applying the synthesis technique to compare the data collected on the SPA process and the yardstick. This comparison would show the weakness points in the evaluated SPA method and be able to suggest improvements in the final evaluation report. The whole evaluation process should be documented. The documentation includes the information about evaluated SPA method, definitions of the evaluation components, data gathered and evaluation results. This documentation would be useful for comparisons with future evaluations of the same or similar targets.

Consequently, the proposed evaluation method is composed of six main components: target; criteria; yardstick; data gathering technique; synthesis technique and evaluation process. The developed evaluation method would be useful in revealing the improvement opportunities of the evaluated SPA method.

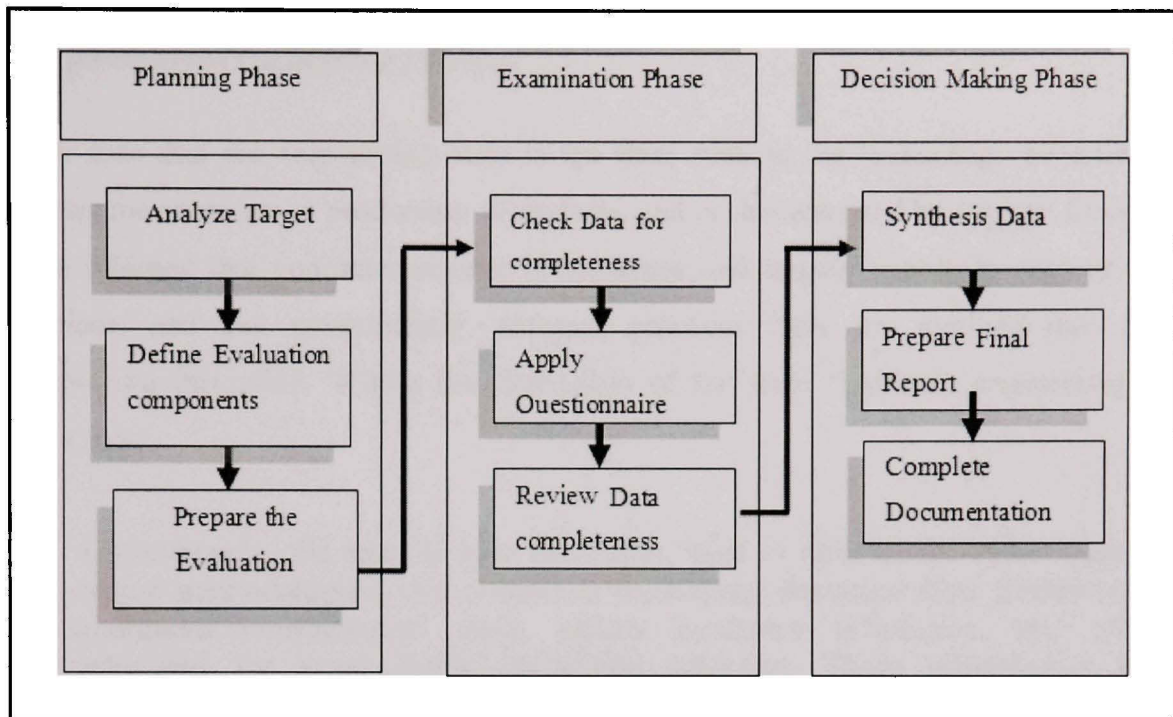


Figure 5.5 Detailed activities for the proposed evaluation process phases.

5.3 Different technological viewpoints

This section presents four different views of technology and the view of software engineering as an engineering discipline.

5.3.1 Shaw's viewpoint

Shaw's view of technology focuses on the supporting sciences (Shaw 1990), rather than engineering itself and has software architecture research as the main example of a supporting field (Shaw 1990; Shaw 2002). (Shaw 1990) presented an evolution model of an engineering discipline as follows:

“Historically, engineering has emerged from ad hoc practice in two stages: First, management and production techniques enable routine production. Later, the problems of routine production stimulate the development of a supporting

science; the mature science eventually merges with established practice to yield professional engineering practice.”

Shaw sees that the engineering state is the final state in the technology evolution which follows the commercial production of artifacts, and is characterized by support from science. Shaw claimed that computer science is immature and largely unable to support software practices, and that consequently, software practices have not evolved into a proper engineering discipline. Within her discussion of the term “software engineering”, Shaw stated that:

“Unfortunately, the term is now most often used to refer to life-cycle models, routine methodologies, cost-estimation techniques, documentation frameworks, configuration-management tools, quality assurance techniques, and other techniques for standardizing production activities. These technologies are characteristic of the commercial stage of evolution - “software management” would be a much more appropriate term”

Therefore, according to Shaw, there are fundamental problems in the use of the term software engineering: “Computer science has contributed some relevant theory, but practice proceeds largely independently of this organized knowledge. Given this track record, there are fundamental problems with the use of the term software engineering” (Shaw 1990).

At the end of her article (Shaw 1990), Shaw discusses five steps to evolve the software profession into a true engineering discipline and focused on the codification of existing knowledge in the form of a handbook such as the Perry’s chemical Engineers’ a handbook, “Simply put, software engineering requires investment in the infrastructure cost - in creating the materials required to organize information, especially reference material for practitioners” (Shaw 1990).

5.3.2 Simon’s viewpoint

The second technological viewpoint introduces the concept of satisfying. A design that works, as an alternative to optimizing, is the best design. This viewpoint considers the design

process as a resource consumer; hence, the designers focus on finding software architecture that works rather than find the optimal architecture. For more details and discussions of this view of technology refer to (Simon 1996).

5.3.3 Vincenti's viewpoint

The third viewpoint of technology considers the technology as a body of knowledge relatively independent from scientific knowledge. Vincenti, in his book "What engineers know and how they know it" (Vincenti 1990), proposed a taxonomy of engineering design knowledge based on the historical analysis of five case studies in aeronautical engineering covering a roughly fifty-year period. He identified different types of engineering design knowledge and classified them into six categories:

1. Fundamental design concepts.
2. Criteria and specifications.
3. Theoretical tools.
4. Quantitative data.
5. Practical considerations.
6. Design instrumentalities.

As stated by Vincenti, this classification is not specific to the aeronautical engineering domain, but can be transferred to other engineering domains. However, Vincenti did not provide documented evidence of this applicability and generalizations to other engineering disciplines.

Vincenti's view of technology as a body of knowledge relatively independent from scientific knowledge implies, according to Jorgensen (Jorgensen 2005), that computer science is an engineering discipline, not a science; Jorgensen mentioned that "If one adopts Vincenti's view of technology as centered on knowledge for design and rooted in its own communities and organizations, it is tempting to suggest that we label as technology or engineering a large

portion of computing disciplines, including portions of information systems and software engineering”. The second part of Jorgensen’s suggestion concerning the software engineering is conformant with Abran et al. in their view of software engineering as engineering discipline as discussed in the next section.

5.3.4 Abran et al. viewpoint

(Shaw 1990) and in her discussion of the steps to be taken to evolve the software profession and software practice into an engineering discipline focused on the codification of existing knowledge in the form of a handbook that is useful for practitioners, The Guide to Software Engineering Body of Knowledge SWEBOK (Abran, Moore et al. 2004) provides a distinguished practical effort in this direction to promote software engineering as an engineering discipline by providing consensus on the core body of knowledge: “For software engineering to be fully known as a legitimate engineering discipline and a recognized profession, consensus on a core body of knowledge is imperative”.

The SWEBOK (Abran, Moore et al. 2004) is subdivided into ten Knowledge Areas - KA as follows:

1. Software requirements.
2. Software design.
3. Software construction.
4. Software testing.
5. Software maintenance.
6. Software configuration management.
7. Software engineering management.
8. Software engineering process.
9. Software engineering tools and methods.
10. Software quality.

The descriptions of the KA are designed to discriminate among the various important concepts, permitting readers to find their way quickly to subjects of interest. Upon finding a subject, readers are referred to key papers or book reports selected because they present the knowledge.

The KA descriptions of software engineering are also forward-looking considering not only what is generally accepted today, but also what could be generally accepted in three to five years. The Guide to SWEBOK was established with the following five objectives:

1. Promote a consistent view of software engineering worldwide.
2. Clarify the place and set the boundary of software engineering with respect to other disciplines such as computer science, project management, computer engineering and mathematics.
3. Characterize the contents of the software engineering discipline.
4. Provide a topical access to the Software Engineering Body of Knowledge.
5. Provide a foundation for curriculum development and individual certification and licensing material.

CHAPTER 6

DEVELOPING EVALUATION METHODS BASED ON ENGINEERING DESIGN PRINCIPLES TO EVALUATE SPA METHODS

6.1 Introduction

In this chapter, the detailed development process for the Top-Down evaluation method is presented. This evaluation method is useful for SPA method designers since it discusses design issues based on engineering design principles. The development of the evaluation method proceeds by defining the six evaluation components. The steps for developing the evaluation method are shown in Figure 6.1.

To be consistent, the evaluation components are presented in sequence even though their definitions were presented in the previous chapter.

6.2 The target

The target for the proposed evaluation is the SPA method. The target concept and its definition in the SPA context have been discussed in section 5.2.1

6.3 The evaluation criteria

In this section, evaluation criteria based on engineering design principles are proposed. Building the evaluation method based on an engineering design viewpoint would help improve the maturity of the SE, mainly the SPA field, as an engineering discipline.

Vincenti in his book (Vincenti 1990) provided a detailed discussion of engineering design knowledge. Vincenti classified engineering design knowledge into six classes. Vincenti's classifications will be used as the bases to build the evaluation criteria for the Top-Down evaluation method.

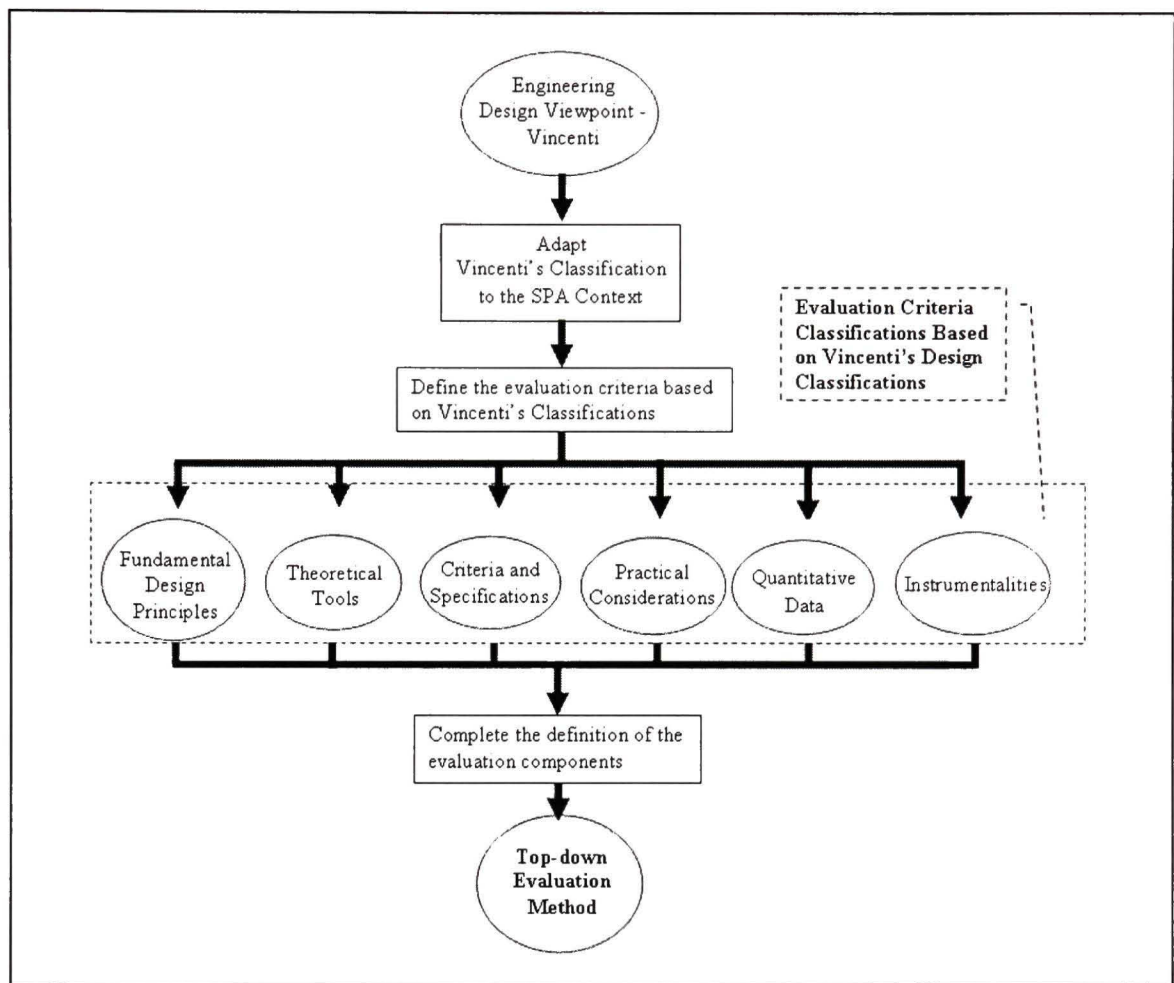


Figure 6.1 Steps to build Top-Down evaluation method.

6.3.1 Mapping Vincenti's classifications to the SPA methods design context

The work done by Vincenti in defining the anatomy of engineering design knowledge based on a long experience in the aeronautical field forms a good framework to study the design process in the SPA field. Vincenti stated that “a complicated technology can often be regarded as a device”. Today, the software products, which are used as stand alone products or as embedded in very complex systems, as well as the development process producing them, are obviously complex technologies and can be regarded as devices performing certain functions. Therefore, Vincenti's classifications are used here to study the SPA methods design from an engineering viewpoint.

Using Vincenti's terms and concepts in this SPA context, designing a new lightweight SPA method for is mostly based on a vicarious model. The common vicarious models used in the SPA field are ISO 15504 and CMMI, which have been adapted to fit the needs of such organizations. Such a vicarious means of selection is preferred as a cost and time saving alternative of building a full assessment model.

In his book, Vincenti discusses the anatomy of design knowledge in the engineering discipline and provides a categorization of engineering design knowledge. This categorization could also be used as an analytical tool to study the coverage of different engineering topics with other domains such as software engineering, for example, the work presented in (Abran and Meridji 2006) proposes some pioneering work in modeling Vincenti's classifications and how to use Vincenti's categories as constituting criteria for investigating software engineering from an engineering perspective.

Accordingly, to investigate the lightweight software process assessment methods using Vincenti's classifications, it is useful to see to what extent the design of these methods aligns with engineering design principles.

Vincenti stated that this classification is not specific to the aeronautical engineering domain only, but can be transferred to other engineering domains. This transfer to the software engineering field in general and software process assessment and improvement in particular, is challenging in the sense that this field is not mature enough to apply such a classification.

As stated by Vincenti, the defined six main classifications, whose breakdown graph is shown in Figure 6.2, are not entirely exclusive since some items of knowledge can embody the characteristics of more than one category. Vincenti also stated that these categories are complete while the details and contents for each of them are not. The contents depend on the domain where this classification is applied. Hence, Vincenti's classifications should be studied in the context of SPA to define the contents relevant to each point in the classification to end up with a form more suitable to be used for evaluating lightweight SPA methods.

The breakdown graph shown in Figure 6.2 would serve as an evaluation criteria tree. In the following sub-sections each of these criteria will be discussed in detail.

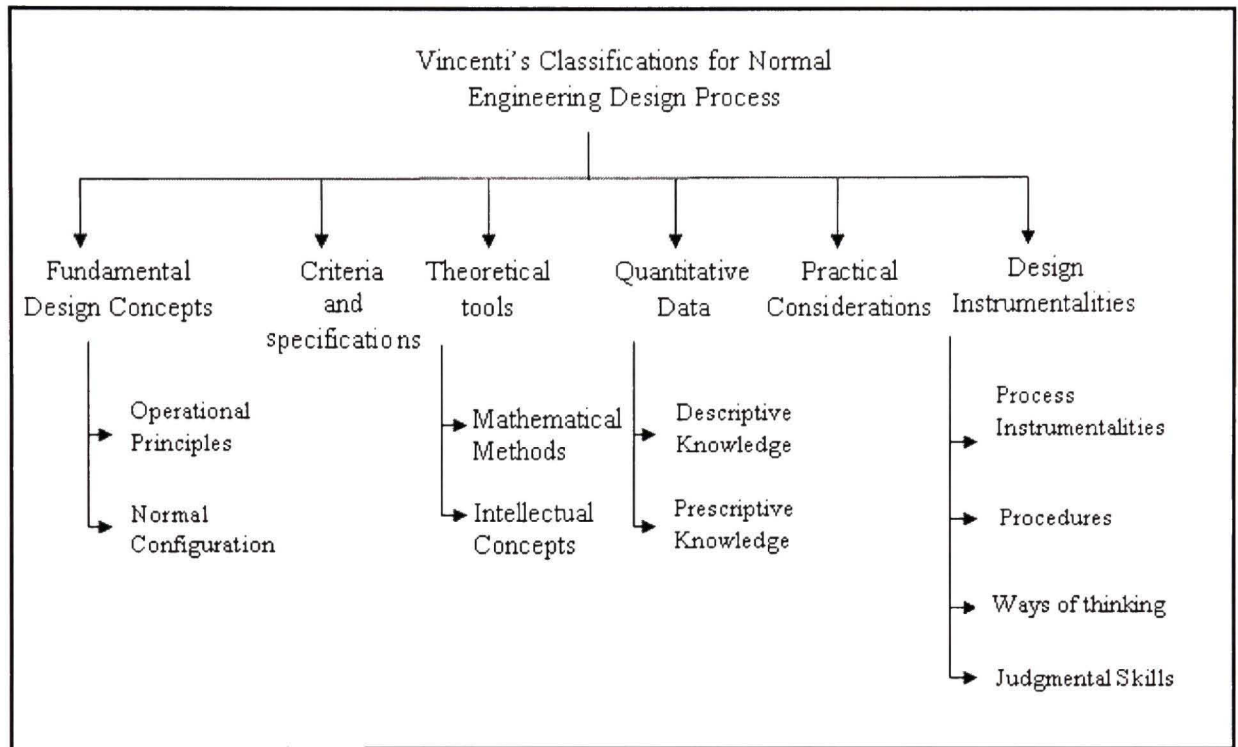


Figure 6.2 Vincenti's classifications breakdown graph – evaluation criteria.

6.3.1.1 Fundamental design principles evaluation criteria

Usually, the designers planning to start a project to build a certain device using a normal design process bring with them some fundamental concepts about the devices. These concepts may exist only in the designers mind implicitly or stated explicitly somewhere else: “they are givens for the projects, even if unstated” (Vincenti 1990). As stated by Vincenti, the fundamental design concepts can be derived from two main sources:

- **Operational principles:**

These principles specify how the different parts of the designed device fulfill special functions in combination with overall operation to achieve the purpose. In other words

“how the device works. The operational principles also, in effect, define a device” (Vincenti 1990).

The main principle to design an SPA method – the proposed device – is that the designer keeps in mind that the software development process should be divided into a set of distinct processes. For each process a clear definition of purpose and outcomes is provided; this is formally known as a process reference model. The designer also keeps in mind that these processes should have indicators which are used to assess the achievement of process attributes. This is formally known as the software assessment model; hence the process reference model and the process assessment model form the main operational principles for SPA methods. Consequently, when evaluating the SPA methods based on the operational principles the following two criteria should be considered:

1. The process reference model which the SPA method is based on.
2. The process assessment model which the SPA method is based on.

- **Normal configuration:**

The normal configuration of a device means “the general shape and arrangement that are commonly agreed to best embody the operational principles”(Vincenti 1990); that is to say, any device or product to be produced, usually, consists of a set of sub-devices or sub-products, the interaction and arrangement of these sub-products is what concerns a normal configuration.

To study the normal configuration in an SPA context, the set of the sub-products and their arrangement should be identified. The best way to do this is by identifying the different phases for conducting an assessment method, and be able to identify the set of sub-products and their arrangements for each phase.

Loon in his book (Loon, Cass et al. 2004) has defined a generic assessment procedure for the assessment process. Despite the fact that this generic procedure is based on an ISO 15504 conformant assessment method for the assessment of the space software processes SPICE for SPACE, this method is general enough to be used for any assessment. The procedure phases are:

- Initiating the assessment;
- Planning the assessment;
- Briefing;
- Data acquisition;
- Data validation;
- Process rating;
- Reporting results.

The details for each of these phases are summarized in Table 6.1. For SME organizations, this generic procedure is still applicable although some details of this procedure can be overlooked for SME organizations. These details include “select the assessment team leader” which gives the impression of having a large assessment team to be lead, whereas the assessment in SME organizations is carried out by one assessor or by an assessor with one assistant assessor only. Another example is related to “select the local assessment coordinator”. Usually such a coordinator is needed when assessing large organizations where the product of one department or unit is interrelated with other departments or units. Hence, to conduct an assessment, the coordination among the participants from these related units should be maintained by the coordinator which is not the case in SMEs: the needed assessment logistics in SMEs can be done via the sponsor or one of the participants without the need to assign an employee to play this role. The cost in time is one of the obstacles facing the SMEs; unnecessary roles should be reduced or eliminated.

As shown in Table 6.1, which summarizes Loon's assessment procedure, the working products proposed by Loon have been divided into two parts: Assessment Input and Assessment Output. The assessment input working products consist of:

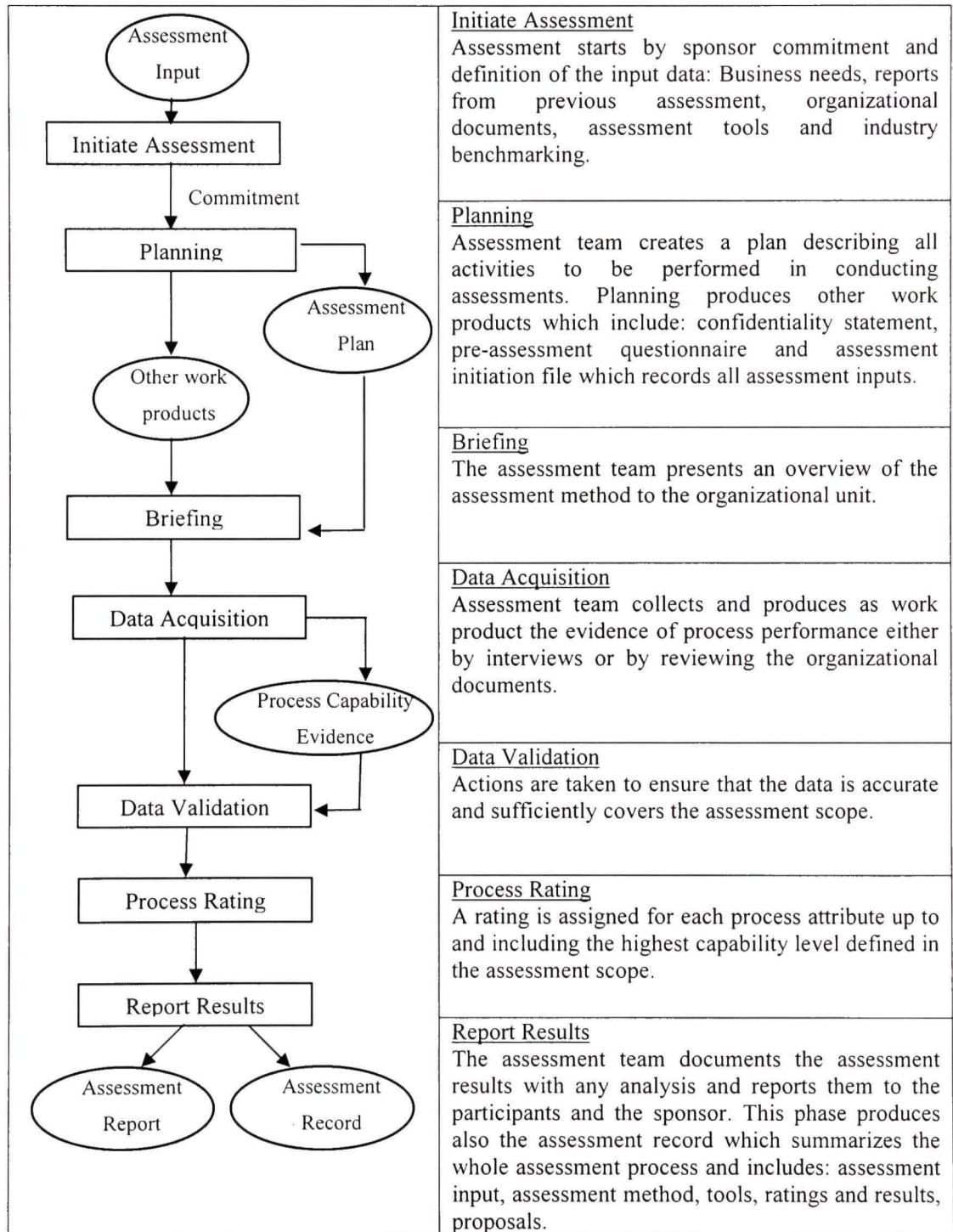
- Business need;
- Reports from previous assessments;
- Organizational documents and reports;
- Assessment instruments.

While the assessment output working products consist of:

- Pre-assessment questionnaire which helps in understanding the organization's goals, software products and the software process currently in use. Due to the lack of software engineering knowledge in SMEs (Wangenheim, Anacleto et al. 2004), revision and completion of the questionnaire through an interview with the representatives of the organization is preferable;
- Assessment initiation file;
- Assessment plan;
- Evidence of process performance and capability;
- Assessment report;
- Assessment record;
- Assessor record.

Table 6.1

Loon's Assessment procedure summary



Consequently, when evaluating the SPA methods based on the configuration management the following criteria should be considered:

1. Define the business need before the assessment.
2. Make use of previous assessment reports.
3. Refer to the organizational documents and reports while preparing for the assessment.
4. Make use of assessment tools through different phases of the assessment.
5. Produce a pre-assessment questionnaire.
6. Produce assessment initiation file.
7. Produce assessment plan.
8. Track evidences of process performance and capability.
9. Produce assessment report.
10. Produce assessment record.
11. Produce assessor record.

6.3.1.2 Criteria and specifications evaluation criteria

Vincenti stated that “to design a device embodying a given operational principle and normal configuration, the designer must have at some point specific requirements in terms of hardware”. When designing a new device, the designer translates the general qualitative goals into specific quantitative goals. The designer must have knowledge of technical criteria appropriate to the device and its use, the designer must also assign numerical values or limits to the characteristics of the appropriate criteria which is essential for the design.

When talking about the criteria and specifications in an SPA context where the device to be produced is a new proposed assessment method, one can define several requirements in terms of the operational principles and the normal configuration incorporated in the design of the intended device.

One of the main models which the SPA process, as mentioned in the operational principles section, is based upon is the software process reference model. The process reference model usually defines the purposes and outcomes of a list of processes. When adopting a certain process reference model to build an assessment method for SME organizations, the designer should decide on the number of processes to be assessed and how they are selected. This selection of processes is important since SMEs are usually interested in some processes but not all of them. Therefore, the following evaluation criteria emerge from this discussion:

1. Specify the number of processes to be assessed.
2. Specify the processes to be assessed.

The other model which the SPA method is based upon is the software assessment model. When choosing the process assessment model to build the new assessment method for SME organizations, the designer keeps in mind that the criteria used to assess each process and what scale is used for measurement and the limits or range of this scale. The designer also keeps in mind the criteria that should be used to assess the organization as a whole if the intent is to assess the whole organization, and keep in mind the scale and its limits to be used to make the measurement. The following evaluation criteria emerge from this discussion:

1. Specify the criteria for assessing the process.
2. Define the scale and its limits used to assess the process.
3. Define the scale and its limits used to assess the organization.

6.3.1.3 Theoretical tools evaluation criteria

Vincenti stated that "To carry out their design function, engineers use a wide range of theoretical tools. These include intellectual concepts for thinking about the design as well as mathematical methods, theories and formulas which can be simple or complex formulas for making design calculations" (Vincenti 1990).

As stated previously, when designing an SPA method the designer should specify the set of processes to be assessed and the mechanism that should be used to define the rate for each process. The designer in order to answer these two questions, uses intellectual concepts to specify which processes to assess and how to rate these processes. Thus, the following evaluation criteria emerge from this discussion:

1. Specify the theoretical tools used to select the processes to be assessed.
2. Specify the theoretical tools used to define the rating process.

6.3.1.4 Quantitative data evaluation criteria

Vincenti focused on the importance of quantities and other data for other physical properties required in the formulas during the design process. Vincenti also stated that “other kinds of data may also be needed to lay out details of the device or to specify manufacturing processes for production” (Vincenti 1990). Such data is usually obtained empirically and sometimes calculated theoretically and are typically represented in tables or graphs. The quantitative data can be divided into two types of knowledge, descriptive and prescriptive (Vincenti 1990).

Descriptive knowledge is the knowledge of how things are. It includes physical constants as well as properties of substances and physical processes. Descriptive data occasionally deal with operational conditions in the physical world.

Prescriptive knowledge is knowledge of how things should be to attain a desired end – it says, in effect, “in order to accomplish this, arrange things this way” (Vincenti 1990).

The quantitative data in the SPA context is greatly related to the rating process. When designing an SPA method, the designer should specify the descriptive data needed to perform ratings either for the process or for the organization. The questions therefore are:

1. What data is used to determine the scale for each process?

2. What data is used to determine the scale for the organization?

6.3.1.5 Practical considerations evaluation criteria

In addition to theoretical tools and quantitative data, Vincenti stated that “Designers also need for their work an array of less sharply defined considerations derived from experience in practice” (Vincenti 1990). Usually, practical considerations are difficult to define and are rarely documented. Sometimes the practical considerations become well codified. In such cases, these practical considerations are moved to another category.

When designing the assessment method for SME organizations, the designers select the set of processes to be assessed either based on their own experience or by applying certain selection methods. When rating the organization with reference to an assessment model, designers also specify the maximum target scaling level to be used based on his experience and according to the needs of the SME organization. The designer also should decide whether to build an action plan or not at the end of the assessment process. Hence, the following criteria emerge from this discussion:

1. How are the processes to be assessed selected?
2. What is the target scaling level for the organization?
3. Does the assessment method build an action plan at the end of the assessment?

6.3.1.6 Instrumentalities evaluation criteria

“Besides the analytical tools, quantitative data and practical considerations required for their tasks, designers need to know how to carry out those tasks” (Vincenti 1990). As part of the engineering design knowledge, the instrumentalities of the design process should be determined which contain the procedures, judgmental skills and ways of thinking by which the process is done.

Vincenti mentioned that “designers doing normal design call upon a number of well-organized, more or less structured procedures”; Vincenti also mentioned that “the division of an overall system into subsystems is fundamental. In the terms of the SPA method, the assessment process is divided into sub-divisions, or phases using Loon’s terms (Loon, Cass et al. 2004). The arrangement and configuration of each sub division is defined by the designer, these sub divisions are executed sequentially when conducting the assessment method which defines the assessment procedure. Examples of such sub-division are the assessment phases defined by Loon and discussed in section 6.2.1.1. These divisions may vary from one assessment method to another and should be evaluated by the evaluation framework. The criterion related to this issue is:

1. Define the sub divisions of the assessment method during the assessment design process.

Another design instrumentality that the designer of a SPA process usually uses are the judgmental skills to define which process should be included in the assessment process; the processes to be assessed are determined based on the organizational objectives and usually, the designer uses his experience and practical considerations to specify these processes. Hence the context on which the assessment is conducted varies from one organization to another and the criterion that should be addressed here is:

2. What are the judgments related to which processes should be taken by the designer?

Vincenti stated that one of the design procedures that can be used to improve the proposed design is the use of iterative techniques, such as successive improvement of a design based on analytical or test experience with earlier versions. Hence, another evaluation criterion emerges:

3. What procedure is used to improve the designed SPA method?

6.4 The yardstick

In this section, the yardstick contents are specified. The yardstick contents shown in Table 6.2 summarize the different yardsticks resulting from the discussion of each of the evaluation criteria which are based on Vincenti's classification. During the design phase of the SPA process, the designer should take these yardsticks into consideration. Failing to take one or more of these yardsticks into consideration would be considered a weak point in the design process that may cause unsuccessful implementation of the SPA process. Consequently, this yardstick has no threshold value or limit. The evaluation would be to verify whether these yardsticks have been considered during the design phase.

Table 6.2

Summary of the SPA method yardstick - Top-Down approach

SPA Process Yardstick Based on Fundamental Design Principles	
No.	Yardstick
SPA-FYS1	Process reference model.
SPA-FYS2	Process assessment model.
SPA-FYS3	Business needs.
SPA-FYS4	The use of previous assessment reports.
SPA-FYS5	Refer to the organizational documents and reports while preparing for the assessment.
SPA-FYS6	The use of assessment tools through different phases of the assessment.
SPA-FYS7	Producing a pre-assessment questionnaire.
SPA-FYS8	Producing assessment initiation file.
SPA-FYS9	Producing assessment plan.
SPA-FYS10	Tracking evidences of process performance and capability.
SPA-FYS11	Producing assessment report.
SPA-FYS12	Producing assessment record.
SPA-FYS13	Producing assessor record.
SPA Process Yardstick Based on Criteria and Specifications	
No.	Yardstick
SPA-CYA1	The number of processes to be assessed.
SPA-CYA2	Processes to be assessed.

Table 6.2

Summary of the SPA method yardstick - Top-Down approach (Continued)

SPA Process Yardstick Based on Criteria and Specifications	
No.	Yardstick
SPA-CYA3	The bases for assessing the process.
SPA-CYA4	Scale and limits used to assess the process.
SPA-CYA5	Scale and limits to assess the organization.
SPA Process Yardstick Based on Theoretical Tools	
No.	Yardstick
SPA-TYS1	Tool used to select the processes to be assessed.
SPA-TYS2	Tool used to define the rating process.
SPA Process Yardstick Based on Quantitative Data	
No.	Yardstick
SPA-QYS1	Data used to determine the scale for each process.
SPA-QYS2	Data used to determine the scale for the organization.
SPA Process Yardstick Based on Practical Considerations	
No.	Yardstick
SPA-PYS1	Selecting the processes to be assessed.
SPA-PYS2	Scaling level for the organization.
SPA-PYS3	Building an action plan.
SPA Process Yardstick Based on Instrumentalities	
No.	Yardstick
SPA-IYS1	The sub divisions of the assessment process.
SPA-IYS2	Designer judgments related to which processes to be assessed.
SPA-IYS3	The procedure used to improve the designed SPA process.

6.5 Data gathering technique

The evaluation data gathering technique can now be formalized as a set of questions each of which deals with one of the requirements mentioned in the yardstick section. The sole work of the SPA method evaluator is to seek an answer for each of these questions. The hypothesis here, as mentioned previously, is that the SPA method that answers as many as possible of these question and takes them into consideration during the design phase, will be more likely

to succeed and hence provide useful and reliable results to be used as an initiator for implementing the suggested improvements.

The questionnaire can be used by designers as a checklist to ensure that all the different criteria have been taken into consideration during the design phase. The questionnaire can also be used by evaluators to collect all necessary information about the SPA method. A sample of the evaluation tool is shown in Table 6.3. The whole evaluation tool is presented in Annex I. The questionnaire consists of a question number; question and answer. The questions are of the ‘open ended’ type.

6.6 The synthesis technique

In the SPA process context, as discussed in section 5.2.5, the multiple values synthesis technique is used where criteria grouping and datum-by-datum comparison with the yardstick is applied.

6.7 The evaluation process

The evaluation process consists of three main phases: planning, examination and decision making. Each of these phases has a set of activities. The main phases and activities associated with each phase are shown in Figure 5.5.

Table 6.3

Sample of the Evaluation Tool - Top-Down approach

SPA Process Evaluation Criteria based on Fundamental Design Principles		
No.	Criteria	Answer
SPA-FQ1	What process reference model is the SPA process based on?	
SPA-FQ2	What process assessment model is the SPA process based on?	
SPA-FQ3	Does the SPA process define the business need before the assessment?	

Table 6.3

Sample of the Evaluation Tool - Top-Down approach (Continued)

SPA Process Evaluation Criteria based on Fundamental Design Principles		
No.	Criteria	Answer
SPA-FQ4	Does the SPA process make use of previous assessment reports?	
SPA-FQ5	Does the SPA method refer to the organizational documents and reports while preparing for the assessment?	
SPA-FQ6	Does the assessment process make use of assessment tools through different phases of the assessment?	

CHAPTER 7

DEVELOPING EVIDENCE-BASED EVALUATION METHODS TO EVALUATE LIGHTWEIGHT SPA METHODS

7.1 Introduction

In this chapter the detailed development process for the bottom-up evaluation method is presented. To be consistent, the evaluation components are presented in sequence while their definitions have been given in chapter 5. To build the bottom-up evaluation method, this chapter looks at what evidences can be used to support the claim of developing a successful SPA method.

An extensive literature review, as discussed in the following section, was conducted to find such evidences, using the guidelines in (Kitchenham 2007) to collect a set of evidences which support the claim of successful implementation of an SPA method. An ‘evidence’ has been defined in (Morse, Swanson et al. 2001) as “an observation, fact, or organized body of information offered to support or justify inferences or beliefs in the demonstration of some proposition or matter at issue”.

This chapter uses the same concept of “evidence” in the SPA context to refer to the published set of requirements, lessons learned, observations and success factors related to the design and implementation of SPA methods. These documented findings will be used, once they are taken into consideration, during the design of the SPA method, as evidences indicate that the SPA method is successful, and would serve as the evaluation criteria for the success of the SPA method. See Figure 7.1.

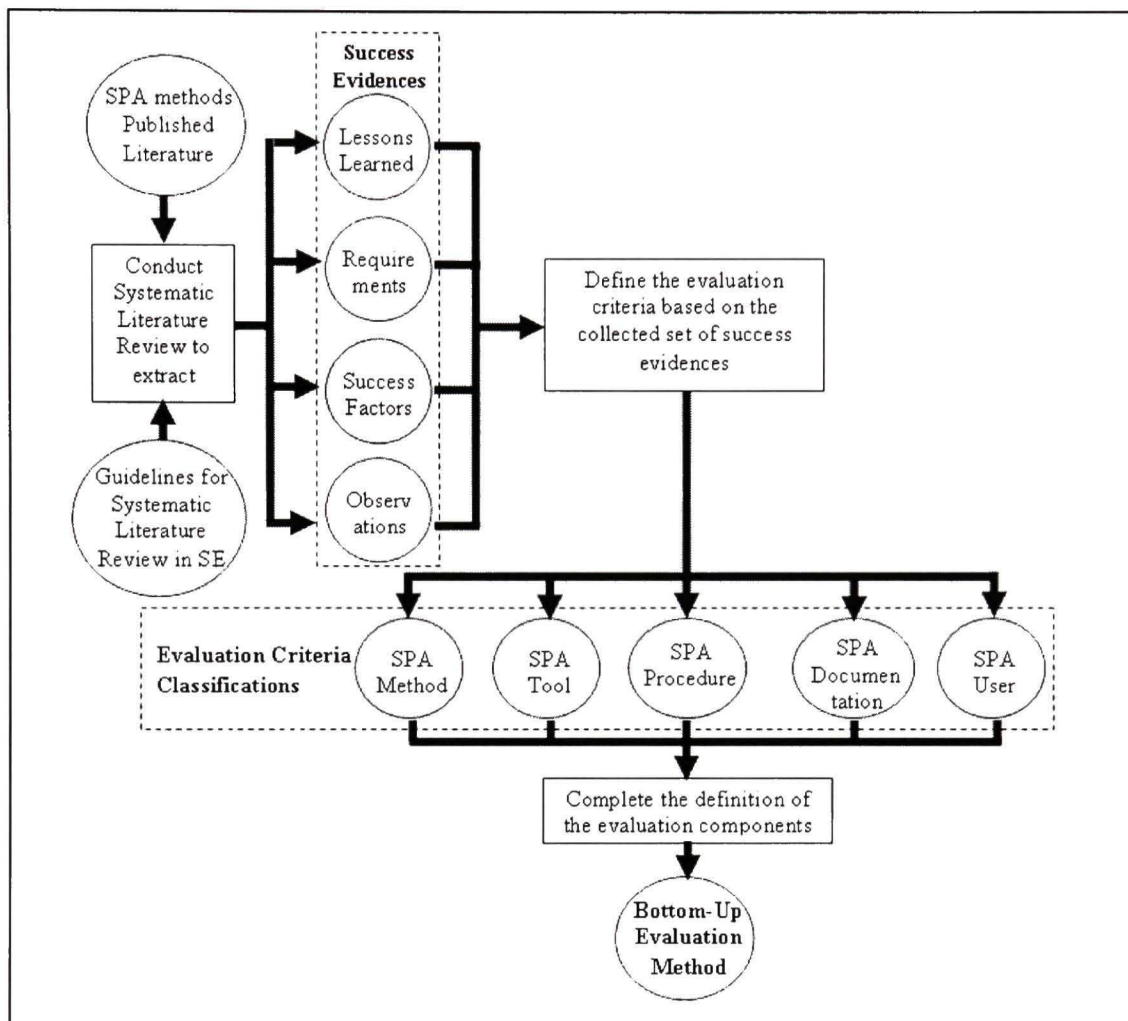


Figure 7.1 Bottom-Up evaluation method development process.

7.2 The systematic literature review

Systematic literature review, also referred to as a systematic review, is a form of secondary study that uses a well-defined methodology to identify, analyze and interpret all available evidence related to a specific research question in a way that is unbiased and (to a degree) repeatable (Kitchenham 2007). This section presents a brief description of the application of the guidelines for performing systematic literature review in software engineering proposed in (Kitchenham 2007) to collect the set of evidences that support the success of the SPA method.

The systematic review guidelines as proposed in (Kitchenham 2007) consist of three main phases as shown in Figure 7.2. These guidelines will be used to in conducting the literature review:

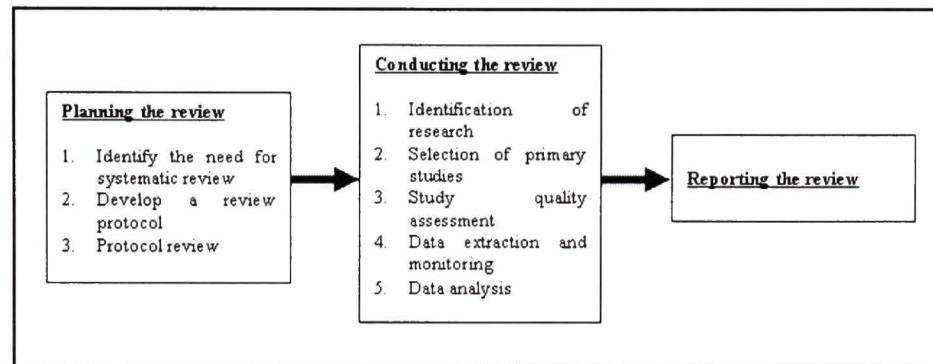


Figure 7.2 Main phases of a systematic literature review.

7.2.1 Planning the review

Planning the review includes the following three activities:

7.2.1.1 Identify the needs for a systematic review

To build an evaluation method that measures the degree of SPA method success, one must collect the evidences for success which are documented in the literature. The literature review tries to answer the following question:

What are the evidences in the literature which can be used to evaluate the success of the SPA method? What are their frequencies?

This need is similar to one listed by (Kitchenham 2007) which is “Assessing the frequency or rate of a project development factor such as the adoption of a technology, or the frequency or rate of project success or failure”. This thesis, firstly, attempts to find what are the evidences affecting the design and implementation success of the SPA method and, secondly, explore which of these evidences are most considered and which are considered least by the

researchers and practitioners in the SPA field by computing the frequencies of these evidences in the documented literature.

Although few studies have been found to conduct a literature review of the success factors affecting SPI programs, no study was found to address the classifications of the success evidences specific to the SPA methods and the building of an evaluation method based on these evidences.

7.2.1.2 Develop the review protocol

An exhaustive search process is conducted using several electronic sources and digital libraries of specific conference proceedings and journal papers. These digital libraries include:

1. EI Compendex (www.engineeringvillage2.org).
2. IEEE Explore (<http://ieeexplore.ieee.org/Xplore/dynhome.jsp>).
3. Science Direct (www.sciencedirect.com).
4. Cite-seer Library (citeseer.ist.psu.edu).
5. Google Scholar (scholar.google.com).

Several search terms are used in conducting the search process in the mentioned resources. These search terms varied from being very simple, such as “software process assessment” to more complex terms such as (“software process assessment” OR “software process improvement”) AND (requirement OR “success factors” OR “lessons learned” OR “observation”).

The search process focused on exploring the articles in the field of software process assessment and improvement published between Jan. 1997 and Dec. 2007 that satisfy the following conditions:

- a. Conducting empirical research through industrial case studies or experiments.
- b. Discuss success evidences of the assessment methods based on the experience gained through the experiments and case studies.
- c. Suggesting possible improvements for the assessment methods.
- d. Collecting and analyzing previous experiences and provide new suggestions and improvements to the assessment methods.

The search process also explored the available technical reports that discuss the requirements and success factors which are directly related to the assessment process regardless of the publication date of these reports.

The articles excluded belong to one or more of the following types:

- a. Articles discussing the software assessment and improvement processes, methods or models without conducting case studies or presenting the author's experience in conducting the assessments.
- b. Articles conducting case studies but without mentioning any assessment related success evidences.

To identify the pieces of data to be extracted from the collected publications and reports, a data model was built as shown in Figure 7.3.

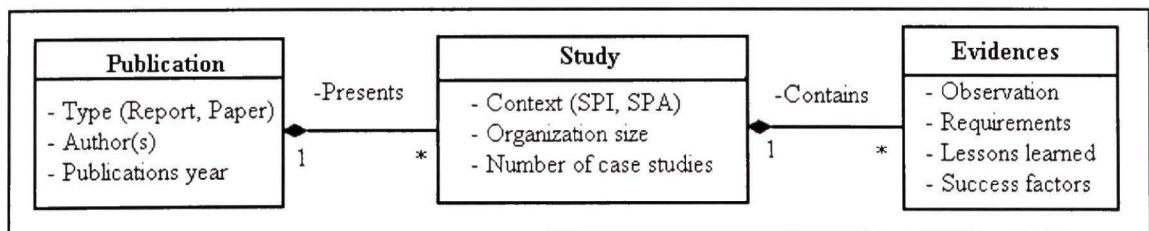


Figure 7.3 Data model for the conducted literature review.

7.2.1.3 Review protocol

As suggested in (Kitchenham 2007) “PhD students should present their protocol to their supervisors for review and criticism”. The review protocol has been verified and approved by the supervisor of this research project.

7.2.2 Conducting the review

Upon the completion of the review protocol, the review process has been conducted as follows:

7.2.2.1 Identification of research

The set of identified publications based on the search terms used have been collected based on their titles. Those documents found relevant and fitting the needs of this study are documented using Microsoft Excel sheet and the format shown in Table 7.1. As can be seen, the table structure includes the different data pieces in the data model, in addition to the classification of these evidences in two more columns: the evidence description as a text and the evidence group ID.

Table 7.1

Archiving format for the set of publications and their evidences

Type of Publication	Publication year	Authors	Context	Organization size
Number of case studies	Evidence description	Evidence Group ID	Evidence	

7.2.2.2 Selection of primary studies

The selection of primary studies went through two main phases: the first phase covered more than 250 published papers and technical reports related to the SPI and SPA field, identified on the basis of their titles and abstracts. In the second phase of the literature review, the papers are read in detail during which the inclusion and exclusion criteria of this research protocol have been taken into consideration; 29 papers were identified as discussing specific evidences related to the design of assessment methods and their application procedures.

7.2.2.3 Study quality assessment

The suggestions of (Kitchenham 2007) to perform a quality assessment of each selected publication is not feasible in this literature review, as this research did not judge the quality of other authors' works and the extent to which those authors identified and controlled validity threats to their work if not already documented. The same idea has been argued by Staples et al. (Staples and Niazi 2007) in their investigation of the systematic literature review guidelines where they mentioned that "We did not feel it would be possible for us (or perhaps any other individuals) to assess the extent to which other authors were able to identify and actually control threats to the validity of their studies". Since this research is in agreement with Staples' viewpoint, no assessment of the quality was made of the collected set of publications. The authors' observations; lessons learned and stated requirements for the SPA process have been considered as candidate evidences for the success of the SPA method.

Although the quality assessment as suggested by (Kitchenham 2007) has been lightly considered in this thesis, the collected set of evidences fit this research's needs.

7.2.2.4 Data extraction

The data have been extracted based on the data model shown in Figure 7.3. The data extracted were the evidences supporting the success of the SPA method. These evidences are

stated explicitly in the selected set of publications, their extraction being a straightforward process. According to the structure shown in Table 7.1, these evidences have been archived along with other extracted information in an Excel sheet and are presented in Annex II.

7.2.2.5 Data analysis

The conducted literature review provides an answer to the research question: “What evidences can be used to support the claim of developing a successful SPA method?” A total of 207 evidences were surveyed and 38 distinct evidences remained after removing duplicates. The set of collected evidences are grouped into five main parts as shown in Figure 7.4.

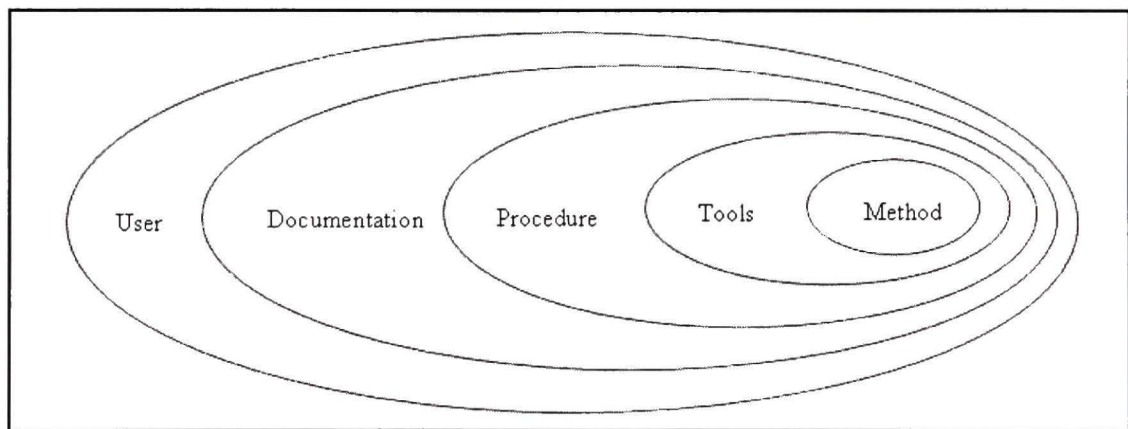


Figure 7.4 Main classes of the collected evidences.

The conducted literature review revealed that the success evidences from conducting an SPI process cover both the assessment phase as well as the improvement phase. Publications that discuss the success evidences are grouped, according to a concept-centric approach (Webster and Watson 2002) into evidences from SPI in organizations of different sizes and evidences from SPI in SME organizations.

1. SPI in organizations of different sizes

The literature review in the SPI field for organizations of different sizes has been conducted to explore the SPA method related success evidences, if any, mentioned as part of the improvement process. The review included published works in which the researchers and practitioners conducted numerous case studies and surveys and where several articles have been found to discuss SPA success evidences - for example: (Lok and Walker 1997; Komiyama, Sunazuka et al. 2001; Fabbrini, Fantini et al. 2003; Chen, Sorenson et al. 2007; Santos, Montoni et al. 2007; Santos, Montoni et al. 2007). The review also covered several technical reports and books that discuss the software process assessment requirements and improvement process - for example: (Humphrey 1989; Loon, Cass et al. 2004; CMMI-SUT 2006).

This literature review revealed that the research perspectives in the field of SPI as a whole, focused in one or more of the following issues:

1. Investigating the critical success factors of the SPI process.
2. Exploring barriers and obstacles to the SPI process.
3. Conducting SPI empirical studies and presenting the experience and lessons learned from such studies.
4. Conducting and reviewing empirical studies of SPA methods to verify the efficiency of these methods. These reviews normally use statistical methods to discuss their findings.
5. Studying the software process improvement frameworks (such as IDEAL model) and how they affect the SPI initiatives.

Most of the research work conducted in the SPI field, as in the review above, focuses on studying the critical success factors (CSF) and barriers facing SPI initiatives, which are based on well-known SPI frameworks, as well as on the lessons learned from conducting case studies. These research initiatives aim at achieving a better understanding of the SPI process in terms of “what” and “how” software process improvements can be achieved.

The identified critical success factors and lessons learned can be divided into two main types, as discussed in (Dangle, Larsen et al. 2005):

- Organizations related CSF such as: leadership involvement, employee participation, commitment, learning ability, experience and cultural changes.
- Process and technology related CSF such as: measures, procedures used, training and reviews.

The classification of such CSFs and the evaluation of the SPI frameworks based on these classifications have been studied by different researchers including (Komi-Sevio 2004). CSFs have also been extensively studied by Niazi who used them to build a maturity model for the implementation of SPI (Niazi, Wilson et al. 2003) (Niazi, Wilson et al. 2005) (Niazi, Wilson et al. 2005) (Niazi, Wilson et al. 2006). An evaluation of the whole SPI process based on CSF is out of the scope of this research project.

Only a few papers and reports have discussed SPA principles and requirements explicitly and in detail. These articles and reports include: (ISO/IEC 2003) (Komiyama, Sunazuka et al. 2001), (Humphery and Kitson 1987), (CMU/SEI 2006) (Loon, Cass et al. 2004) (Wangenheim, Varkoi et al. 2006) and Fabbrini (Fabbrini, Fantini et al. 2003) who proposed an assessment tool and discussed the assessment requirements in general as well as the requirements for an assessment supporting tool. These documents will be discussed in detail later on.

In summary, the field of SPI has been mostly studied in the past two decades in terms of the SPI models created, SPI methods applied and the numerous number of empirical studies conducted to understand better the SPI process, as well as the factors and barriers affecting success. Such empirical studies have been conducted in different organizations of varying sizes: large, medium, small and very small organizations.

Researchers in the field of SPI for SME organizations have noticed that the SPI models used mainly by large organizations, which are usually comprehensive models, are difficult to use in SME organizations, as documented in most of their publications: for example:

- Laryd and Orci (Laryd and Orci 2000) mentioned that “there is a need for models for small organizations. The existing models for software process improvement, e.g. CMM, are overkill for several reasons, and more over difficult to understand and comprehend for the management of small organizations”;
- Johnson and Brodman (Johnson and Brodman 1997) stated that “The CMM largely reflects the software practices of large businesses and large software organizations; moreover, many of its practices are inappropriate to small projects, which are prevalent not only in small businesses and small software organizations but also in large businesses”;
- Kelly and Culleton (Kelly and Culleton 1999) mentioned that “smaller organizations often operate under different constraints” compared to large organizations;
- Laporte and April (Laporte and April 2005) stated that “ISO international standards were not written for small projects, small development organizations, or companies with between 1 and 25 employees, and are consequently difficult to apply in such settings”;
- Similarly, Villalón and his colleagues (Villalón, Cuevas et al. 2002) stated that “Current software process improvement methods (i.e. ISO 15504, CBA-IPI,...) are difficult to apply to small and medium-sized enterprises SMES due to the cost e.g. financial, time and recourses associated with their application”.

These comments suggest that the barriers facing the SPI process have much more influence on SME organizations compared to other types of organizations and seem to be related to the structural, organizational and managerial nature of SME organizations.

These differences have motivated some researchers and practitioners to investigate and design new assessment and improvement models and methods to meet the needs of the SME organizations. These new models and methods are usually tailored from, and conformant to,

the models and methods used by other types of organizations. As a result of this trend in the SPI field, several research and empirical studies are now focusing on the SPI process in SME organizations, as illustrated in the next section.

2. SPI in SME and VSE organizations

The second part of the literature review focused on the papers that discussed the software process assessment and improvement in SME organizations, mainly, those papers that presented the success factors, lessons learned or observations on the SPI process in SME organizations, as discussed in section 7.1.1.2. Examples of such papers include: (Habra, Eustache Niyitugabira et al. 1999; Habra, Niyitugabira et al. 1999; Wiegers and Sturzenberger 2000; Cater-Steel 2001; Salvaneschi, Grasso et al. 2006; Stambollian, Habra et al. 2006; Pettersson, Ivarsson et al. 2008).

As the literature review above shows, researchers have created new SPI models and methods aimed at SME organizations associated also with a number of empirical studies. Although the researchers in this part of SPI field claimed that their new models and methods are conformant to other well-known models and methods, little work has been done to validate or evaluate such assertions, see (Cater-Steel 2004).

This thesis is focused on building an evaluation method to assist in evaluating the different SPA methods used by SME organizations as a starting phase for their SPI initiative. The evaluation method developed in this research thesis could also be used by researchers planning to build new SPA method to ensure that their method fulfills the requirements of process assessment methods.

So far, the study of success in the SPI field discusses mainly the improvement process. Assessment is mentioned implicitly in such studies and hence no assessment specific issues are mentioned when discussing SPI success. Therefore, the research work presented in this thesis complements the work already done in this field aimed at ensuring a successful SPI

process by studying the assessment process separately to see what makes assessment methods design and implementation succeed.

Less attention is paid to the success evidences of lightweight SPA methods, and to what extent these evidences are taken into consideration in the current and proposed lightweight SPA method. The hypothesis in this research thesis is that an explicit evaluation of the lightweight SPA requirements would help in providing a successful implementation of lightweight SPA methods, and hence participate in conducting a successful SPI process as a whole.

7.2.2.6 Reporting the review

The result of the systematic review is not organized in a separate report as suggested by (Kitchenham 2007) as it will not be presented externally. The results of this literature review will be used as input to define the evaluation criteria of the evaluation method proposed in this thesis.

7.3 Detailed discussion of SPA success evidences

In this section, a sample of publications which include books, technical reports and published articles that addressed SPA methods success evidences in the form of requirements, observations, success factors and lessons learned are discussed. A summary of the publications that have been reviewed in this thesis and found to discuss specific success evidences related to the SPA method are presented in Annex II. As shown in Figure 7.3 section 7.1.2.5, the collected set of evidences is grouped into five main classes, namely, SPA methods, supportive tools, procedures, documentation and users. The list of publications supporting each these classes is presented in section 7.2.8. The documents presented in this section are discussed in detail one by one in an author-centric way and are summarized in Table 7.2.

The set of collected and classified evidences and the proposed method will be applicable to any lightweight SPA method since the evidences for successful SPA method are general and applicable for all methods. This claim conforms with the viewpoint stated in (Komiya, Sunazuka et al. 2001) where the authors concluded that “the same requirements for conducting successful assessments are common to all SPA methods”.

Table 7.2

Sample of the documents discussing SPA success evidences

Reference	Description	Type
(ISO/IEC 2003)	ISO/IEC 15504-3 Information technology - Process Assessment — Part 3: Guidance on performing an assessment.	ISO standard
(CMU/SEI 2006)	Appraisal Requirements for CMM1, V1.2.	Technical Report
(Humphery and Kitson 1987)	Preliminary Report on Conducting SEI-Assisted Assessments of SE Capability.	Technical Report
(Loon, Cass et al. 2004)	Process Assessment and Improvement.	Book
(Wangenheim, Varkoi et al. 2006)	Standard Based SPA in small Companies.	Journal Paper
(Komiya, Sunazuka et al. 2001)	Proposal on Library-Centered SPA.	Journal Paper
(Fabbrini, Fantini et al. 2003)	Performing SPICE assessment: yet another tool.	Conference Paper

7.3.1 ISO/IEC 15504-3 (ISO/IEC 2003): guidance on performing an assessment

This part of ISO/IEC 15504 provides guidance on meeting the minimum set of requirements for performing an assessment contained in ISO/IEC 15504-2 and specifies a set of factors that are essential to a successful process assessment. These factors are:

1. Commitment:

ISO/IEC 15504-3 stated that “The commitment of the sponsor is essential to ensuring that the assessment objectives are met. This commitment requires that the necessary resources, time and personnel are available to perform the assessment”.

2. Motivation:

ISO/IEC 15504-3 states that the organization's management “needs to motivate participants to be open and constructive. Process assessments focus on the process, not on the performance of Organizational Unit members implementing the process. The intent is to make the processes more effective in support of the defined business goals, not to allocate blame to individuals”. This part of the standard also states that “Providing feedback and maintaining an atmosphere that encourages open discussion about preliminary findings during the assessment helps to ensure that the assessment output is meaningful to the Organizational Unit”.

3. Confidentiality:

ISO/IEC 15504 -3 mentions that “Respect for the confidentiality of the sources of information and documentation gathered during assessment is essential in order to secure that information. Where interviews or discussions are employed, consideration should be given to ensuring that participants do not feel threatened or have any concerns regarding confidentiality”.

4. Relevance :

ISO/IEC 15504-3 mentions that “The Organizational Unit members should believe that the assessment will result in some benefits that will accrue to them directly or indirectly”.

5. Credibility:

ISO/IEC 15504 states that “The sponsor, the management and the staff of the Organizational Unit should all believe that the assessment will deliver a result which is objective and is representative of the assessment scope. It is important that all parties can be confident that the assessors have adequate assessment experience, are sufficiently impartial and have an adequate understanding of the Organizational Unit and its business to conduct the assessment”.

ISO 15504-3 provides a description and guidance for the assessment requirements defined in part 2 of the same standard (ISO/IEC 2003). These requirements “aim at achieving a greater degree of uniformity in the approach to process assessment, so as to maximize the reliability of different approaches and provide a degree of comparability between the results of different assessments” (ISO/IEC 2003). The requirements for performing assessments as described in the ISO 15504-3 (ISO/IEC 2003) are summarized in Annex IV.

7.3.2 SCAMPI requirements (CMU/SEI 2006)

The SCAMPI appraisal team has published a document that defines the Appraisal Requirements for CMMI – ARC - which are considered essential for the appraisal methods based on CMMI. The ARC document has been designed “to help improve consistency across multiple disciplines and appraisal methods and to help appraisal developers, sponsors and users understand the tradeoffs associated with various methods” (CMU/SEI 2006). The appraisal methods have been classified into three main classes to provide guidance for the developers to specify the class of the applications most suitable for their appraisal methods. Hence, the appraisal methods are defined to be either class A, B or C. The requirements are then assigned to each class based on the attributes of this class.

The differentiation between the three classes, as stated in the report, is based on attributes such as (CMU/SEI 2006):

1. The degree of confidence in the appraisal outcomes.
2. The generation of ratings.
3. Appraisal cost and duration.

Class A methods must satisfy all the ARC requirements and can be used to provide ratings for benchmarking. The developers of such methods have the option to conduct ISO 15504 conformant appraisals. Class B appraisal methods are required to comply with a subset of the

ARC requirements. Class B methods do not produce ratings and are not intended to be ISO 15504 conformant. These types of appraisals are recommended for initial assessments in organizations that are just beginning to use CMMI models for process improvement activities. Class C appraisal methods are required to comply with a subset of the ARC requirements for Class B methods. Validation and corroboration are also optional for Class C methods. These types of appraisals would most likely be used when the need for a “quick look” arises or for periodic self-assessments by projects and organizational support groups.

Based on the descriptions above for classes A, B and C; class C is obviously the most appropriate to develop lightweight SPA methods since the other two classes would result in large assessment methods which require considerable effort from both the assessed organizations as well as the assessor. As noticed by McCaffery (McCaffery, McFall et al. 2005) who has devised with his colleagues an assessment method based on CMMI for small organizations and argues that the class C method is very suitable for small and medium size organizations. McCaffery mentioned that his method “complies with the ARC 1.1 requirements for a class-C method” (McCaffery, McFall et al. 2005).

Accordingly, the set of requirements for methods of class C as discussed in (CMMI-SUT 2006), and presented in Annex III, have been studied as the best candidate requirement for developing lightweight SPA methods. For simplicity, the optional requirements for class C methods have not been considered.

7.3.3 Watts Humphrey (Humphrey and Kitson 1987)

Humphrey in his technical report (Humphrey and Kitson 1987) described the SEI methodology for conducting SEI-assisted assessments of software engineering capability. When implementing an assessment method (Humphrey and Kitson 1987) stated that “There are several critical prerequisites to a successful assessment which must be thoroughly understood and accepted by the assessment participants”. These prerequisites are:

1. Confidentiality:

Humphrey states that that “Confidentiality is required at all levels of the organization. All the professionals who talk must be told that what they say will not be attributed to them”.

2. Senior management involvement:

The senior manager should be involved personally in the assessment process since this person must be convinced of the action’s importance if anything is to happen.

3. Non-adversarial attitude:

- The assessment process should be non-threatening and focus on learning from local professionals and understanding the current organization status;
- The assessment process should tap the knowledge and creative skills of the best local people, to help improve the organization;
- A highly critical attitude or lack of interest in local views and opinions can be disastrous.

4. Action orientation:

The entire motivation of the assessment must be directed toward improvement. The orientation is on actions, so the questions must focus on defining those problems that need to be solved right away, otherwise the assessment may make the problems worse.

7.3.4 Van Han Loon (Loon, Cass et al. 2004)

Loon has mentioned five factors which he considers as essential to successful process assessment, as also mentioned in ISO 15504-3 (ISO/IEC 2003). These factors are:

1. Commitment:

Loon states that “commitment to the complete assessment process and use of the results is of vital importance”. The sponsor must be committed to the assessment purpose and provide the authority and resources to undertake the assessment within the organization.

2. Motivation of participants:

The management unit of the organization should motivate the participants to be open and constructive. The participants should be aware that the assessment process focuses on the process not on the members implementing it.

3. Confidentiality:

The participants in the assessment from the organization side are the principle source of knowledge and experience about the process. Respecting the confidentiality of this source of information during the assessment is essential to secure the information. The participants should not feel threatened in order to provide the knowledge and experience they have regarding the process.

4. Benefits:

As stated by Loon: “The organizational unit members should believe that the assessment will result in some benefits that will accrue to them directly or indirectly”.

5. Credibility:

Loon states also that “the sponsor and the management and staff of the organizational unit must all believe that the assessment will deliver a result, which is objective and representative of the assessment scope”.

7.3.5 Christiane Gresse Wangenheim (Wangenheim, Varkoi et al. 2006)

In (Wangenheim, Varkoi et al. 2006), Wangenheim et al. have studied the application of standards, ISO 15504, in software process assessment for small organizations. Several lessons learned point to the importance of some issues to the assessment process:

1. Flexibility of assessment model based on a continuous representation.
2. Focus the assessment process on the principle high-priority processes.
3. Coverage of the process reference model.
4. Data collection based on group interviews.
5. Identification of risks and improvement opportunities.
6. Availability of documents templates and tool support.

7.3.6 Toshihiro Komiyama (Komiyama, Sunazuka et al. 2001)

(Komiyama, Sunazuka et al. 2001) proposed a framework for software process assessment and improvement which is adaptable to the features of the assessed organization. The proposed framework is based on the experience of the authors on the application of different SPA methods at NEC; the motivation for their proposed framework is that, although they applied several SPA methods, they were still unable to decide which method was the best. In their work, Komiyama and his colleagues discussed several SPA requirements which are classified into three main parts:

1. Assessment procedure

1.1. Self-Assessment

- 1.1.1. Specific questions should be developed prior to the interview to obtain consistently interpreted process features; if not, the assessment data will not be reliable.
- 1.1.2. There should not be too many questions in order to reduce the assessment workload.

1.2. On-Site Assessment

- 1.2.1. Reserve a minimum of two hours for interviews to obtain detailed process status, but do not go beyond one full day of questioning.
- 1.2.2. Prioritize the questions in order of importance, especially if there are many questions.
- 1.2.3. A surveillance and interview of general project's status during on-site assessment is helpful to obtain the information to prioritize questions.
- 1.2.4. Before interviewing, review documentation such as development plan, progress reports and specifications to flag any interview questions.

1.3. Reporting

- 1.3.1. Assessment report is composed of two parts: one part can be drafted systematically which contains rating results and produced graphs and the other part describes the improvement proposal.

- 1.3.2. Support tools to analyze and visualize the assessment data are useful for producing the first part of the report.

2. Assessment method

- 2.1. There is no best SPA method. Methods should be easy to customize for each organization's goals, needs and properties.
- 2.2. SPA methods should be usable for both self-assessment and on-site assessment. Also, it is desirable that the collected data be compatible between methods.
- 2.3. The number of questions for one interview should be no more than 150, ideally less than 100. The wider coverage and the finer granularity the better, but may lead to a greater number of questions. Balance is necessary.
- 2.4. A well-structured questionnaire makes finding correlations easier and gets answers with fewer questions. A well-structured questionnaire will reduce the interview time.
- 2.5. The role or position of the interviewee should be clarified to get reliable and correct answers.
- 2.6. A roadmap along with milestones for process improvement should be provided that prioritizes the established issues.
- 2.7. A means to indicate the effects of process assessment and improvement quantitatively and objectively should be provided.
- 2.8. The relationship between product quality or project results and process quality should be clarified based on analysis of project data and assessment results. Prioritizing the process improvement actions to be taken is helpful.

3. Assessment tool

The required functions for the supporting tools are:

- 3.1. Assessment data collection and analysis.
- 3.2. On-line assessment support.
- 3.3. Analysis and visualization of assessment data.
- 3.4. Database of historical SPA data.
- 3.5. (Semi-) Automatic assessment-report generation.
- 3.6. Library of knowledge and experience on process assessment and improvement.

Komiyama concluded that the urgent and important requirements are summarized into three main requirements:

1. Adaptability - The possibility to adapt SPA methods to organizational needs, goals, and properties.
2. Concreteness - The possibility to reach effective and concrete solutions based on assessment results.
3. Validity - The possibility to validate the effects of process assessment and improvement activities.

7.3.7 Fabbrini (Fabbrini, Fantini et al. 2003)

(Fabbrini, Fantini et al. 2003) discussed, while presenting their new assessment tool, the requirements for 'SPICE assessment' to satisfy the needs of different stakeholders, such as the assessors, sponsors and organization participants. The authors also discussed the requirements that an automatic tool should support the different phases of the assessment. These requirements are as follows:

- **The assessment requirements:**

1. Reliability - One can rely on assessment result to allow organization unit (OU) management to take correct decision about the OU goals.
2. Objectiveness and Repeatability - Results do not depend on the particular assessment team and/or assessment candidate.
3. Completeness - Any essential element to assess each process in the assessment scope is taken into account and all the needed results are given.
4. Documentability - All the normative items and other aspects, such as assessors' decisions regarding information source selection and process attribute ratings are reported in a way that is easy to retrieve.
5. Cost-effectiveness - Assessment tools should balance the cost and results obtained.

- **Assessment supporting tool requirement**

1. **Assessment Coverage** - The assessment tool should be able to support as many as possible of the assessment activities.
2. **Usability** - The assessor would like to use the tool during different phases of the assessment.
3. **Adaptability/Flexibility** - Assessments can differ in terms of the application domain to which the process instance belongs, in terms of size or in terms of assessment duration. Because the assessment may depend on these characteristics, the tool has to be able to be adapted to them.
4. **Meaningfulness** - The information provided by the tool must be correct, updated, and compliant with the model, in general, be able to provide the assessor with real help.
5. **Inclusiveness** - A tool supporting an assessment should include enough information to increase the confidence of having taken into account all the relevant aspects of process instance.
6. **Ability to interact with the expert assessor.**

7.3.8 Literature review summary

In this section, several publications are discussed which address some of the success evidences for SPA methods design and implementation. According to the conducted systematic review, 29 publications have been found to discuss the SPA methods success evidences explicitly and in detail. Table 7.3 summarizes the publications relating to the main classes of the collected set of success evidences. A total of 207 success evidences have been collected. These evidences have been filtered into 38 distinct success evidences grouped into five main classes. The extracted information from each of the selected references is shown in Annex V.

Table 7.3

Publications relating the five classes of collected success evidences

Success evidence class	Publications
SPA method related success evidences.	(CMMI-SUT 2006), (Humphrey 1989), (Wangenheim, Varkoi et al. 2006), (Anacleto, Wangenheim et al. 2004), (Anacleto, Wangenheim et al. 2004), (Fabbrini, Fantini et al. 2003), (Kautz 1998), (McCaffery, Taylor et al. 2007), (Komiyama, Sunazuka et al. 2001), (Grünbacher 1997), (Stambollian, Habra et al. 2006), (Laporte, Desharnais et al. 2005), (Habra, Eustache Niyitugabira et al. 1999), (Dyba and Moe 1999), (Wiegers and Sturzenberger 2000), (ISO/IEC 2003), (Salvaneschi, Grasso et al. 2006),
SPA supportive tool related success evidences.	(Wangenheim, Varkoi et al. 2006), (Anacleto, Wangenheim et al. 2004), (Anacleto, Wangenheim et al. 2004), (Fabbrini, Fantini et al. 2003), (Komiyama, Sunazuka et al. 2001), (Ekdahl and Larsson 2006), (Wiegers and Sturzenberger 2000), (ISO/IEC 2003), (Choi, Lee et al. 2005), (Chen, Sorenson et al. 2007), (Lok and Walker 1997),
SPA procedure related success evidences.	(Loon, Cass et al. 2004), (CMMI-SUT 2006), (Humphrey 1989), (Anacleto, Wangenheim et al. 2004), (Cater-Steel, Toleman et al. 2006), (Cater-Steel 2002), (Fabbrini, Fantini et al. 2003), (Komiyama, Sunazuka et al. 2001), (Grünbacher 1997), (Stambollian, Habra et al. 2006), (Laporte, Desharnais et al. 2005), (Dyba and Moe 1999), (Ekdahl and Larsson 2006), (Wiegers and Sturzenberger 2000) , (Santos, Montoni et al. 2007), (Santos, Montoni et al. 2007)

Table 7.3

Publications relating the five classes of collected success evidences (Continued)

Success evidence class	Publications
SPA documentation related success evidences.	(CMMI-SUT 2006), (Wangenheim, Varkoi et al. 2005), (Anacleto, Wangenheim et al. 2004), (Anacleto, Wangenheim et al. 2004), (Cater-Steel, Toleman et al. 2006), (Fabbrini, Fantini et al. 2003), (Ekdahl and Larsson 2006), (Wieggers and Sturzenberger 2000), (ISO/IEC 2003),
SPA user related success evidences.	(Loon, Cass et al. 2004), (CMMI-SUT 2006), (Humphrey 1989), (Anacleto, Wangenheim et al. 2004), (Cater-Steel 2002), (Komiyama, Sunazuka et al. 2001), (Grünbacher 1997), (Ekdahl and Larsson 2006), (Wieggers and Sturzenberger 2000), (ISO/IEC 2003), (Cater-Steel 2001), (Salvaneschi, Grasso et al. 2006), (Pettersson, Ivarsson et al. 2008), (Santos, Montoni et al. 2007), (Santos, Montoni et al. 2007)

It is important to mention here that the bottom-up evaluation method defines five classes of evidences and, for each class, a set of related evidences are defined. Hence, the developed evaluation method has two levels: the classes of evidences and the evidences themselves for each class. A third level of evidences can be added to the developed bottom-up evaluation method. The bases of the third level of evidences will be discussed in the “future work” section.

Figure 7.4 clarifies the developed part of the bottom-up evaluation method and the new level which can be developed in the future.

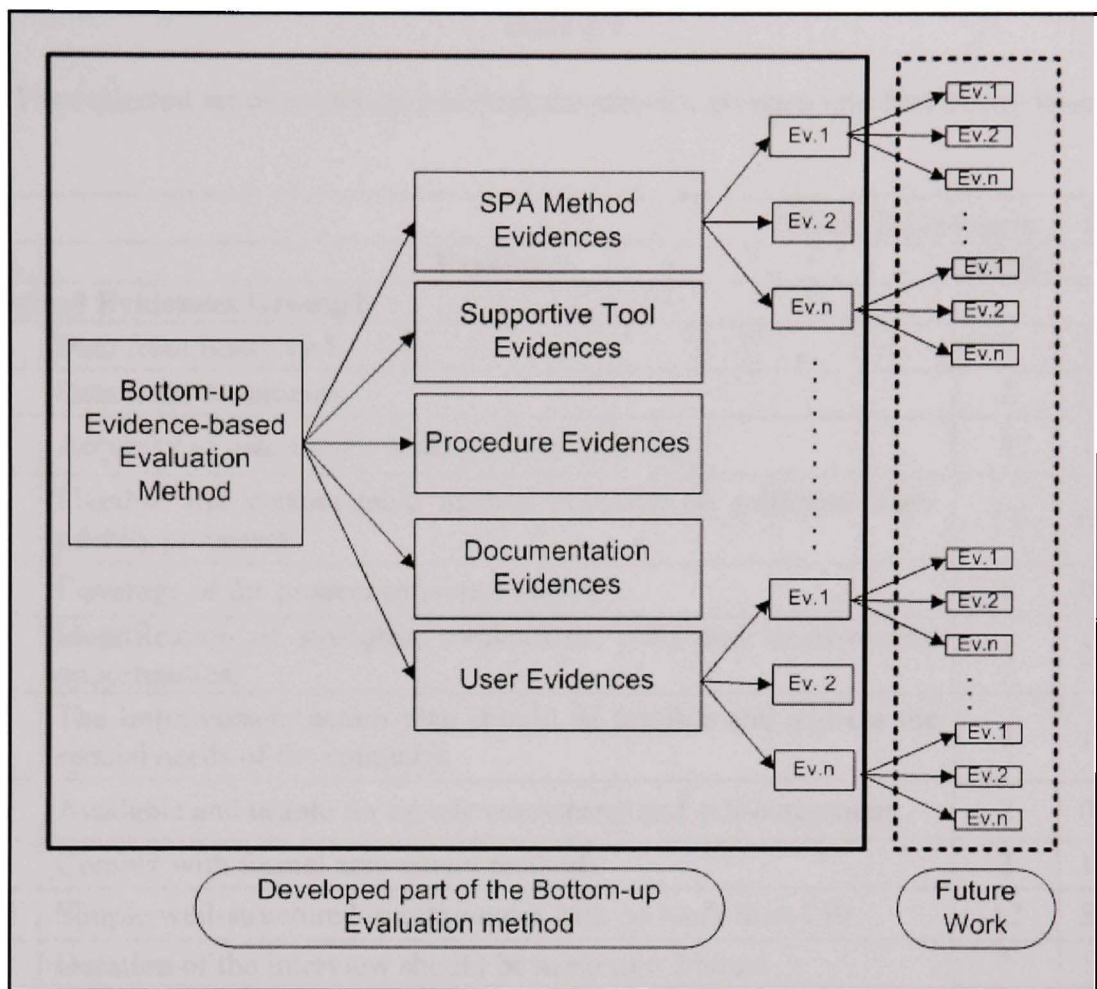


Figure 7.5 Extending the developed Bottom-Up evaluation method by adding a third level.

The collected set of evidences is shown in Table 7.4 along with their frequencies and groupings into the defined five main classes. The frequencies of different success evidences, as presented in Figure 7.6, show that the research community in the SPA field has different levels of concerns with the different classes of success evidences.

Table 7.4

The collected set of evidences and their frequencies grouped into five main classes

Total of Evidences			207
No.	Evidence	Freq.	%
Method Evidences Group-2			
1	Data from interviews.	3	1.45
2	Data from documents.	2	0.97
3	Accuracy of assessment findings (data collected).	3	1.45
4	Flexible and customizable method focusing on principal high-priority processes.	13	6.28
5	Coverage of the process reference model.	1	0.48
6	Identification of strengths, weaknesses, risks and improvement opportunities.	5	2.42
7	The improvement action plan should be feasible and address the special needs of the company.	4	1.93
8	Available and usable for on-site assessment and self-assessment.	2	0.97
9	Comply with formal assessment method.	3	1.45
10	Simple well-structured questionnaire with no more than 150.	12	5.80
11	Duration of the interview should be minimum 2 hours.	3	1.45
12	Reliability and repeatability of the assessment result.	6	2.90
13	Completeness.	1	0.48
Tool Evidences Group-3			
1	Usable support tools which cover the different phases of the assessment process including collect, analyse and visualize the assessment data.	19	9.18
2	Build and use database of historical SPA data.	5	2.42
3	(Semi-) Automatic assessment-report generation.	1	0.48
4	Adaptability/Flexibility.	2	0.97
5	Support confidentiality of assessment by providing a tool that has well-defined security features, roles and permissions.	2	0.97
6	Ensure repeatability of results.	1	0.48

Table 7.4

The collected set of evidences and their frequencies grouped into five main classes
(Continued)

Total of Evidences			207
No.	Evidence	Freq.	%
Procedure Evidences Group-4			
1	Preparation of the assessment process (e.g. introductory sessions, training, define assessment input, develop assessment plan).	25	12.08
2	Build confidence and trust relationships with sponsors and assessment participants.	4	1.93
3	Produce assessment report delivered to the organization.	3	1.45
4	Ensure confidentiality.	6	2.90
5	Hold feedback session after each assessment and a follow-up session (if applicable) after improvement.	8	3.86
Documentation Evidences Group-1			
1	Guidance for identifying assessment purpose, objectives and logistics.	2	0.97
2	Guidance for identifying organizational unit.	1	0.48
3	Guidance for assessment team (assessment team background, credentials, and responsibilities).	7	3.38
4	Guidance for ensuring confidentiality.	2	0.97
5	Providing document templates.	4	1.93
6	Documentation of the assessment method and its implementation in practice (including assessment plan, initiation files, assessment record, assessor record, any discussions during the assessment).	16	7.73
7	Documentation of results of data collection and ratings.	5	2.42
8	Guidance to the follow-up assessors.	1	0.48
User Evidences Group-5			
1	Organization participants responsibilities.	6	2.90
2	Assessment team credentials and responsibilities.	5	2.42
3	Senior management and other staff members Involvement.	7	3.38
4	Commitment.	8	3.86
5	Benefits: the participants should feel the benefits of the assessment.	5	2.42
6	Credibility: the sponsor and staff should believe that the assessment will give a result.	4	1.93

According to these statistics, one can identify the success evidences classes that have been weakly addressed by the researchers and practitioners in the SPA field that could be a source of failure of the SPA methods implementation.

For example, as seen in Table 7.5, the success evidences related to the SPA method itself have gained the largest interest of the researchers and practitioners in the SPA field with around 28% of the collected evidences. The evidences related to assessment procedure come in the second place with around 22% of the collected set of evidences followed by the evidences related to documentation with around 18%, followed by the evidences related to users with around 17%, and in the last place, come the evidences related to supportive tools with around 15%.

Table 7.5

Statistics summary of the success evidences found in the literature

Evidence Class	Freq.	%
Method Evidences.	58	28.02%
Supportive Tool Evidences.	30	14.49%
Documentation Evidences.	38	18.36%
Procedures Evidences.	46	22.22%
User Evidences.	35	16.91%
Total	207	100.00%

7.4 Evaluation components definitions

This section presents the definition of the evaluation components in detail:

7.4.1 The target

The target for the proposed evaluation is the SPA method. The target concept and its definition in a SPA context has been discussed in section 5.2.1.

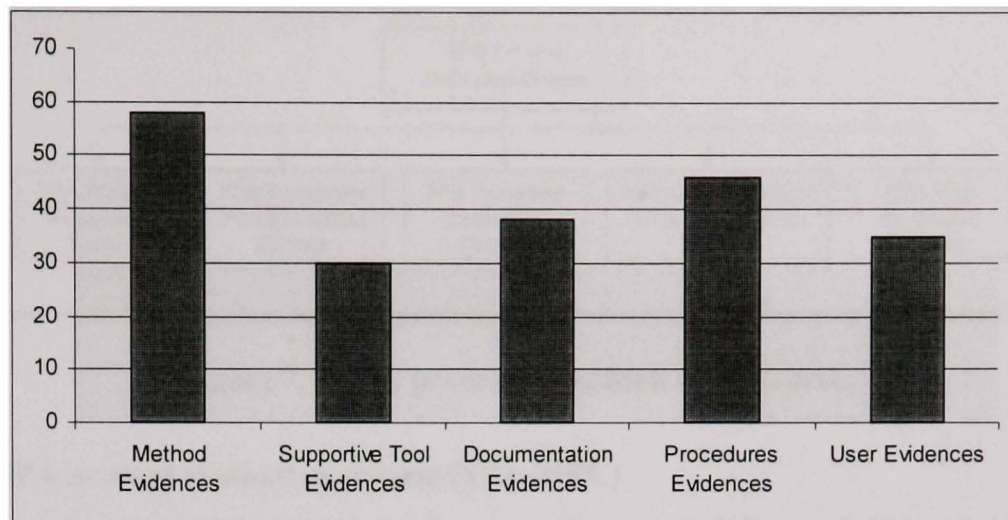


Figure 7.6 Research community interests in the different classes of success evidences.

7.4.2 The evaluation criteria

The evaluation criteria in the proposed evaluation method are based on the collected set of evidences explored from the literature review, as discussed in the previous sections. The explored evidences are those found necessary by researchers and practitioners to implement a successful SPA method.

Consequently, and based on the classified set of evidences, the evaluation criteria for SPA methods are divided into five main parts as shown in Figure 7.7:

1. Method Evaluation Criteria.
2. Supportive Tool Evaluation Criteria.
3. Procedure Evaluation Criteria.
4. Documentation Evaluation Criteria.
5. User Evaluation Criteria.

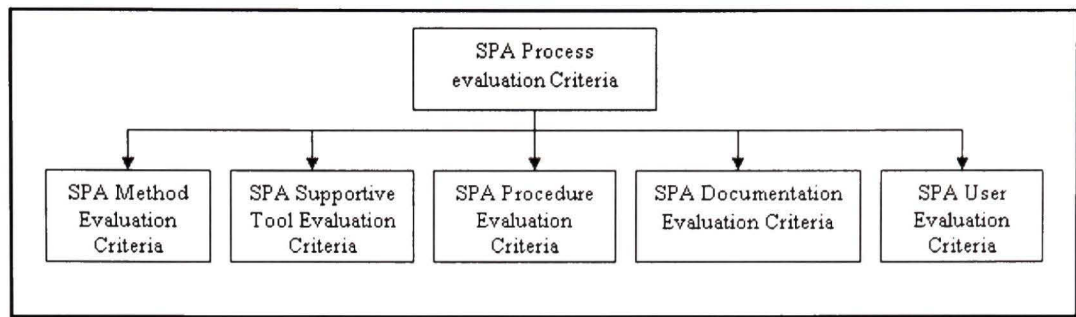


Figure 7.7 SPA process evaluation criteria tree.

7.4.2.1 SPA method evaluation criteria (SPA-MEC)

The SPA method evaluation criteria are based on the evidences which have been discussed by several researchers in the SPA field and include:

- (Humphrey and Kitson 1987) discussed several critical prerequisites to successful assessment;
- (Wangenheim, Varkoi et al. 2006) mentioned several lessons learned at the end of their experiments for their ISO standards based assessment methods;
- (Komiya, Sunazuka et al. 2001) discussed several requirements for assessment methods, procedures and tools;
- (Fabbrini, Fantini et al. 2003) in their presentation of a new assessment tool discussed several requirements for both the assessment method and the supportive tools;
- (McCaffery, Taylor et al. 2007) mentioned several organizations' demands which were taken into consideration in their assessment method as an advantage of their proposed method to help organizations achieve their goals;
- (Anacleto, Wangenheim et al. 2004) at the end of their experiments in applying an assessment method dedicated for small companies based ISO 15504 discussed several observations and critical success factors regarding the applied assessment method;
- (Grünbacher 1997) discussed several guiding principles for his proposed assessment process;

- (Kautz 1998) in his discussion of how and under what circumstances software process improvement can be rewarding concluded that one of the success factor is the use of “a flexible, tailored assessment and improvement approach”;
- (Richardson 2002) discussed several characteristics for self-assessment improvement model. These characteristics are useful and vital to design the assessment method. The list of collected requirements is filtered to remove redundancy. These requirements and their definitions are specified in the yardstick section.

7.4.2.2 SPA tool evaluation criteria (SPA-TEC)

The importance of the availability of support tools for the assessment process has been discussed and mentioned as a requirement for SPA processes by several researchers: (Wangenheim, Varkoi et al. 2006) (Komiyama, Sunazuka et al. 2001) (Fabbrini, Fantini et al. 2003) (Anacleto, Wangenheim et al. 2004) (Anacleto, Wangenheim et al. 2004) (Grünbacher 1997) and (McCaffery, Taylor et al. 2007). This importance has been indicated in these references either as a requirement, lesson learned, support tool function or as observations. The collected and filtered set of evidences is specified in the yardstick section.

7.4.2.3 SPA procedure evaluation criteria (SPA-PEC)

The requirements of the procedure section are aimed at focusing on the way the assessment is conducted and the necessary preparations. Several researchers (Komiyama, Sunazuka et al. 2001) (Grünbacher 1997) (Humphery and Kitson 1987) (Anacleto, Wangenheim et al. 2004) (Cater-Steel, Toleman et al. 2006) and technical reports such as (CMU/SEI 2006) have mentioned some requirements and improvements to the way the assessment is conducted based on their experience. Others (McCaffery, Taylor et al. 2007) mentioned the organizations’ demands as another requirement (“the time to prepare and perform the assessment should be minimal”). The collected set of findings from these references are filtered and grouped into a set of requirements specified in the yardstick section.

7.4.2.4 SPA documentation evaluation criteria (SPA-DEC)

Several researchers focused on the importance of the documentation of the SPA process in their discussions of the experience they gained in the SPA process for small and very small organizations. In presenting how the assessments based on standards are conducted, (Wangenheim, Varkoi et al. 2006) focused on the documentation as a lesson learned: “The assessment effort can be considerably reduced when templates for the documents to be produced during an assessment are available”. The detailed description of the assessment process, the definition of the assessment methods and the guidance for process selection have been emphasized by (Anacleto, Wangenheim et al. 2004). Further improvement for the SPA process can be achieved by providing and documenting the guidance to follow-up meetings as mentioned by (Cater-Steel, Toleman et al. 2006). The CMMI technical report (CMU/SEI 2006) has documented numerous requirements needed for the SCAMPI appraisal method including the version dedicated to small organizations.

7.4.2.5 SPA user evaluation criteria (SPA-UEC)

This section has been written based on several requirements and success factors mentioned by several researchers: (Humphrey and Kitson 1987; Komiyama, Sunazuka et al. 2001; Loon, Cass et al. 2004) and Technical report (CMU/SEI 2006). These evaluation criteria focus on several issues such as specifying the responsibilities for both the assessment team members and the participants from the organization side and give confidence to the assessment process and ensure commitment of the participants in the assessment process

7.4.3 The yardstick

In this section the yardstick content which is built based on the evaluation criteria is specified.

7.4.3.1 Yardstick based on SPA method

Table 7.6

SPA method yardstick (SPA-MYS)

SPA method yardstick	
No.	Yardstick
SPA-MYS1	Data collection from interviews.
SPA-MYS2	Data from documents.
SPA-MYS3	Accuracy of assessment findings (data collected).
SPA-MYS4	Flexible and customizable method focusing on high priority processes.
SPA-MYS5	Coverage of the process reference model.
SPA-MYS6	Identification of strengths, weaknesses, risks and improvement opportunities.
SPA-MYS7	Suggesting improvement action plan.
SPA-MYS8	Available and usable for on-site assessment and self-assessment.
SPA-MYS9	Comply with formal assessment method.
SPA-MYS10	Simple well-structured questionnaire with no more than 150 question.
SPA-MYS11	The assessment duration should take between 2-8 hours.
SPA-MYS12	Reliability.
SPA-MYS13	Completeness.

- **SPA method yardstick description**

- SPA-MYS1: Data collection from interviews:

Since one of the goals of the assessment is to collect assessment data, the assessment method should include interviews with the participants from the organization side, such as managers and developers.

- SPA-MYS2: Data collection from documents:

The collected assessment data can also be gathered by reviewing the documents available in the organization. These documents, when available, can provide data useful for the assessment. The assessment method can use either the interview or the documents review to collect the assessment data.

- SPA-MYS3: Accuracy of assessment findings (data collected):

The collected data should be consolidated into accurate findings according to defined criteria. The following criteria have been stated by (CMU/SEI 2006) to consolidate the collected data into accurate findings:

- a. “The finding was derived from objective evidence seen or heard during data collection sessions”;
- b. “The finding is clearly worded, phrased without attribution, and expressed in terminology used at the organizational unit”;
- c. “Objective evidence supporting the finding is traceable to the project or organizational unit”;
- d. “The finding is relevant to the appraisal reference model and can be associated with a specific model component”.

- SPA-MYS4: Flexible and customizable method focusing on high priority processes:

This requirement states the need for a mechanism to help selecting the processes to be assessed for each organization based on its business goals and needs. As mentioned in (Humphery and Kitson 1987) “the entire motivation for the assessment must be directed toward improvements. The orientation is action”; this means that the assessment method should focus on finding the serious problems and high priority process areas to be improved that are most important for the organization.

- SPA-MYS5: Coverage of the process reference model:

The assessment method in its selection of processes should support a process reference model. The assessment method should also support as many as possible of the activities mentioned in the process reference model for each selected process.

- SPA-MYS6: Identification of strengths, weaknesses, risks and improvement opportunities:
As one of the goals of the assessment process, the assessment must define the strength and weaknesses of the assessed processes as well as the improvement opportunities and any possible risks.
- SPA-MYS7: Suggesting an improvement action plan:
At the end of the assessment and considering the results produced based on SPA-MYS7, an improvement action plan should be defined to draw the roadmap for the proposed improvements.
- SPA-MYS8: Available and usable for on-site and self assessment:
The assessment method should be usable for both self assessment conducted by the organization itself as well as on-site assessment conducted by an external assessor.
- SPA-MYS9: Comply with a formal assessment method:
Lightweight assessment methods are usually tailored from more comprehensive assessment methods. Compliance with comprehensive assessment method enhances the success opportunities of the tailored assessment methods.
- SPA-MYS10: Simple method having a well-structured questionnaire with no more than 150 questions:
This requirement indicates the need for simple balanced assessment methods that take into consideration the length of the assessment method in terms of questions, which should not be more than 150 (typically less than 100) and the coverage of all details for the assessed processes: “The wider coverage and the finer granularity, the better but if a greater number of questions is needed, they should be balanced” (Komiya, Sunazuka et al. 2001).

The question should be organized in a well-structured questionnaire to reduce the cost of the assessment time and help the assessor to be organized during the interview to get results for each question clearly.

- SPA-MYS11: assessment duration take between 2-8 hours:

The assessment duration should be as short as 2 hours with a maximum of 8 hours.

- SPA-MYS12: Reliability:

As stated by (Fabbrini, Fantini et al. 2003), reliability of the assessment method means that “one can rely on assessment results to allow the OU which undertakes the assessment process to take correct consequent decisions about the OU goals”, where OU means the organization unit involved in the assessment process.

Reliability of the assessment method will also produce repeatable results. Different assessment initiatives conducted by different assessment teams for the same processes and same conditions should get the same results.

- SPA-MYS13: Completeness:

As stated by (Fabbrini, Fantini et al. 2003) completeness means “any essential element to assess each process in the assessment scope is taken into account and all needed results are given”.

7.4.3.2 Yardstick based on SPA support tool

Table 7.7

SPA supportive tool yardstick (SPA-TYS)

SPA Supportive Tool Yardstick	
No.	Yardstick
SPA-TYS1	Build and use a database of historical SPA data.

Table 7.7

SPA supportive tool yardstick (SPA-TYS) (Continued)

SPA Supportive Tool Yardstick	
No.	Yardstick
SPA-TYS2	Support different phases of the assessment process including collect, analyze and visualize assessment data.
SPA-TYS3	Generation of semi-automatic assessment report.
SPA-TYS4	Adaptability / Flexibility.
SPA-TYS5	Support Confidentiality of assessment.
SPA-TYS6	Ensure repeatability of results.

- **Support tool yardstick description**

- SPA-TYS1: Build and use a database of historical SPA data:

Building a database after the assessment process which contains the process profiles and other necessary data would be useful for new assessment trials and also for comparing assessment results with previous assessment trials.

- SPA-TYS2: Supporting different phases of the assessment process including collect, analyze and visualize assessment data:

The assessment support tool should provide help to the assessor during the assessment process for all assessment activities and should be able to collect assessment data during the assessment process. The tool should then make analyses based on the underlying assessment model and visualize the data, providing the assessor and the organization with meaningful data about the conducted assessment.

- SPA-TYS3: Generation of a semi-automatic assessment report:

For a successful and more efficient assessment process, one recommends that the assessment supportive tool be able to semi-automate part of the final documents produced by the assessment process.

- SPA-TYS4: Adaptability / Flexibility:
The assessment tool should be able to adapt to different situations where the assessment process may change in terms of duration or application domain of the assessed process instance as mentioned by (Fabbrini, Fantini et al. 2003).
- SPA-TYS5: Support Confidentiality of assessment:
The assessment tool should provide security features that support the obligations of the confidentiality agreement with the organization. The security features may include preventing unauthorized access, using passwords, auditing and other features.
- SPA-TYS6: Ensure repeatability of results:
The assessment tool should provide repeatable results during the rating process each time the assessment is conducted.

7.4.3.3 Yardstick based on SPA procedure

Table 7.8

SPA procedure yardstick (SPA-PYS)

SPA Procedure Yardstick	
NO.	Yardstick
SPA-PYS1	Preparing the assessment process.
SPA-PYS2	Building confidence and trust relationship with sponsors.
SPA-PYS3	Produce assessment report delivered to the company.
SPA-PYS4	Ensure Confidentiality.
SPA-PYS5	Hold feedback session held after each assessment.

• Procedure yardstick description

- SPA-PYS1: Preparing the assessment process:
The preparation to conduct an assessment method includes preparing the participants in the assessment process. The preparation phase includes:
 - a. Provide necessary training for the organizations participants.

- b. Develop an assessment plan specifying the assessment activities, schedules, necessary resources and other issues that may affect the assessment process.
- c. Produce an initiation file which includes all the inputs for the assessment.

The aim of this preparation is to create a sufficient consensus of the proposed assessment including:

- a. Clarifying the purpose, scope and method of the assessment.
- b. The roles and responsibilities of the participants.
- c. Confidentiality of conducted assessments.
- d. Proposed schedules and activities of the assessment.

- SPA-PYS2: Building confidence and trust relationships with sponsors:

A vital condition for the success of the assessment method is to build bridges of confidence and trust between the sponsors and assessors through face to face meetings instead of phone calls or emails, to discuss and plan for the assessment process before the assessment.

- SPA-PYS 3: Produce an assessment report delivered to the organization:

The produced assessment reports along with all the details of the assessment method and its implementation should be provided to the sponsors and the assessed organization. This report will help the organization take decisions regarding improvement actions to be conducted later on. The report includes:

- a. The process profile.
- b. Weakness and strengths related to the assessed process.
- c. Improvement possibilities.
- d. Possible risks.

- SPA-PYS 4: Ensure Confidentiality:

The assessment results must be kept in strict confidence. No leaks can occur, even to the organization's chief executive. Senior management has a proper interest in the results;

but if the members of the organization learn that they cannot really speak in confidence, trust will be lost that will be nearly impossible to rebuild. Confidentiality is required at all levels of the organization. All the professionals who talk must be told that what they say will not be attributed to them. The assessors should provide certain procedures that ensure the confidentiality of the assessment input provided by the participants.

- SPA-PYS 4: Hold feedback session after each assessment:

After conducting the assessment method, the assessor should hold a feedback session to present the results of the assessment to the assessed organization and to discuss the participants' comments and suggestions; these discussions can enable continuous improvement of the assessment method.

7.4.3.4 Yardstick based on SPA documentation

Table 7.9

SPA documentation yardstick (SPA-DYS)

SPA Documentation Yardstick	
No.	Yardstick
SPA-DYS1	Guidance for identifying assessment purpose, objectives and logistics.
SPA-DYS2	Guidance for identifying organization unit.
SPA-DYS3	Guidance for assessment team.
SPA-DYS4	Guidance for ensuring confidentiality.
SPA-DYS5	Documentation of the document templates.
SPA-DYS6	Documentation of the assessment process and its application in practice.
SPA-DYS7	Documentation of results of data collection and ratings.
SPA-DYS8	Guidance for the follow-up assessors.

- **Documentation yardstick description**

- SPA-DYS1: Guidance for identifying assessment purpose, objectives and needed resources:

Providing detailed information about the purpose of the assessment to be conducted, the objectives and recourses needed for each assessed organization must be documented.

- SPA-DYS2: Documentation of the guidance for identifying an assessed organization unit:
As part of the documentation of the SPA process, the organization unit to be assessed should be defined; this process includes defining the current projects of this unit, the participants from this unit in the assessment activities and any other related information.
- SPA-DYS3: Documentation of the guidance for the assessment team:
This part of documentation defines the assessment team members' background, experience as well as their levels of knowledge in the assessment method and underlying model. The documentation also defines the team leader's management and technical skills related to the assessment method.
- SPA-DYS4: Documentation of the confidentiality agreement concerning the assessment data:
The confidentiality of the assessment data is considered important for a successful SPA method; the interviewees will feel safer, be more comfortable and be forthcoming when they know that the information they give in the assessment will not be of any threat to them.
- SPA-DYS5: Documentation of the document templates:
The availability of templates for the documents to be produced at the end of the assessment would help in reducing the effort of reporting the assessment results.
- SPA-DYS6: Documentation of the assessment process:
The entire assessment process should be documented and includes:
 - a. Description of the assessment process.
 - b. Guidance for process selection.
 - c. Definition of the assessment method and underlying model.
 - d. Assessment plan and initiation file.
 - e. Assessment record and assessor record.
 - f. Any discussions during the assessment.

- SPA-DYS7: Documentation of data collection results and ratings:

The set of collected data and the resulting ratings should be documented and added to the assessment report. This process can be done with the help of the assessment supportive tool and would help reduce the cost of the assessment process.

- SPA-DYS8: Documentation of the guidance for follow-up meetings:

Providing guidance for the assessors who will conduct the follow-up meetings, by providing a procedure for the follow-up meeting, would help improve the assessment results. This guidance should be documented at the end of the assessment method implementation.

7.4.3.5 Yardstick based on SPA users

Table 7.10

SPA user yardstick (UYS)

SPA User Yardstick	
No.	Yardstick
SPA-UYS1	Organization participants' responsibilities.
SPA-UYS2	Assessment team responsibilities.
SPA-UYS3	Senior management involvement.
SPA-UYS4	Commitment of participants.
SPA-UYS5	Assessment participant believe that the assessment will give results.
SPA-UYS6	Assessment participants feel the benefits of conducting an assessment.

- **User yardstick description**

- SPA-UYS1: Defining the assessment participants' responsibilities:

The responsibilities of the assessment sponsor and interviewees should be defined. The assessment sponsor should verify that the assessment will provide the expected results by verifying the skills and experience of the assessment team and verifying the confidentiality of interviewee input.

- SPA-UY52: Defining the assessment team responsibilities:

The assessment method should document the responsibilities of the assessment team leader (usually there is one assessor) which include: ensure that the assessment is conducted in accordance with the underlying assessment method; ensure the commitment of the sponsors; ensure the availability of necessary documents for the assessment; ensure that the assessment requirements are met; ensure the readiness of any participating assessors.

- SPA-UY53: Senior management involvement in the assessment process:

Senior management should be committed to the implementation of the assessment method and be involved by attending the meetings, participating in setting priorities for assessed processes and improvement actions.

- SPA-UY54: Ensuring commitment by the participants:

The assessment sponsors should be committed to the implementation of the assessment method. The assessment team leader responsibility, as mentioned earlier, is to ensure the sponsor's commitment.

- SPA-UY55: The assessment team believes the assessment will give results:

Another important aspect that would encourage effective involvement in implementing the assessment method, mainly by interviewees, is the belief that the assessment will give results and will improve the organization's behaviour and performance.

- SPA-UY56: The assessment participants feel the benefits of assessment:

The assessment team should feel the benefits of conducting an assessment on their organization and the effects of that assessment in the improvement process later on. The assessment team should work on building such assurance during the preparation phase of the assessment. Being confident of the benefits would promote better involvement and cooperation on the part of the organization participants in the assessment process.

7.4.4 Data gathering technique

The data gathering technique can now be formalized as a set of questions, each of which deals with one of the evidences mentioned in the yardstick section. The task of the SPA method evaluator is to record an answer for each of these questions. The hypothesis here, as mentioned previously, is that the SPA method which addresses fully, as many as possible, of these question (i.e. takes them into consideration during the design phase) will meet most, if not all, of the success evidences and provide useful and reliable results to be used for initiating the implementation of any suggested improvements.

The questionnaire developed in this research thesis, as an evaluation tool, can be used by the designers as a checklist to ensure that all the different criteria have been taken into consideration during their design phase. The questionnaire can also be used by the evaluators to collect all necessary information about the evaluated SPA method. An evaluation tool representing the questionnaire has been also developed as an Excel sheet.

A sample of the evaluation tool is presented in Table 7.11. The whole evaluation tool is presented in Annex VI. The questionnaire consists of a question number, question, answer and comments; the answer for the questions is a scale of three values where F is Fully Adequate, P is Partially Adequate and N is Not Adequate. The comments column is used by the evaluator to provide explanations supporting or explain the answer.

7.4.5 The synthesis technique

In an SPA context, as discussed in section 5.2.5, the multiple values synthesis technique is used where the datum-by-datum comparison with the yardstick is applied to see to what extent each of the yardsticks applies. Each yardstick should be fully adequate to ensure successful implementation of the SPA process.

Table 7.11

Sample of the evaluation tool - Bottom-Up approach

SPA Method Evaluation			
No.	Question	Answer	Comments
SPA-MQ1	Does the method acquire assessment data from interviews?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ2	Does the method acquire assessment data from documents?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ3	Does the method ensure the accuracy of findings	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ4	Is the method flexible and customizable (i.e. possibility of adding new axes) by focusing on high priority processes using certain mechanism?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ5	Does the method provide coverage to a process reference model?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ6	Does the method identify strengths, weaknesses, risks and improvement opportunities?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	

7.4.6 The evaluation process

The evaluation process consists of three main phases: Planning, Examination and Decision making. Each of these phases has a set of activities. The main phases and activities associated with each phase are shown in Figure 5.5.

CHAPTER 8

RESEARCH APPROACH VERIFICATION PROCESS

8.1 Introduction

This chapter verifies the research process adopted in this thesis, using a set of verification strategies. “Strategies of verification are those techniques that contribute to the validity of the project and are implemented in the actual research process” (Morse, Swanson et al. 2001). The applied verification techniques are based on information from within the research project. “Many of the within-project techniques to ensure verification of findings are standard methods used to control the threats to validity and to ensure reliability” (Morse, Swanson et al. 2001).

8.2 Type of research conducted

The research conducted in this thesis is, mainly, qualitative research. The qualitative research approach is usually used for the investigation of social phenomena and data produced are represented as words/text and pictures, rather than numbers (Gilgun 1992). Adapting qualitative methods into the design of empirical studies in software engineering has been studied by different researchers, see (Seaman 1999).

The main focus of this research work is, on one hand, to align the design of SPA methods with engineering design principles. This alignment required exploration of the classifications of the engineering design process and using these classifications as bases to design SPA methods. On the other hand, this research focuses on exploring and understanding the researchers’ and practitioners’ experiences in designing and implementing successful lightweight SPA methods.

Accordingly, two evaluation methods have been proposed. Each of these methods uses a questionnaire/interview to collect data on the evaluated SPA method. The type of the data collected is words/text. To be able to present the results and conduct comparisons among different SPA methods, the resulting qualitative data is transformed into quantitative data through a coding process. In the coding process, the evaluation participants' responses are mapped on an ordinal scale (Fully given value 1.0, Partially given value 0.5 and No given value 0.0).

8.3 Verification techniques

Verification is the process of reviewing, confirming, making sure, and being certain. In qualitative research, verification refers to the mechanisms used during the process of research to incrementally contribute to ensuring reliability and validity (Morse, Barrett et al. 2002).

The verification techniques discussed in (Morse, Swanson et al. 2001) were used as the basis for this verification process. These techniques include:

a) **Situating the project: the literature review:**

Qualitative research is used “when little is known about a topic” (Morse, Swanson et al. 2001). “This merely means that a study being targeted in the planned project has not been conducted, that few related studies were identified in the literature search, or that they are somewhat scarce” (Morse, Swanson et al. 2001). In the context of SPA methods design and implementation, this research project has been situated according to the following bases:

- From an engineering design viewpoint, the literature review revealed that the alignment of SPA method design with engineering design principles, based on Vincenti's classifications, has not been studied before;
- The conducted literature review revealed that no formal evaluation of lightweight SPA methods has been conducted so far. Formal evaluation is based on evaluation

theory concepts. The literature review showed that the major work in this field is focused on comparisons among different SPA methods, as discussed in chapter 4.

b) Project design:

“Study design constitutes the plan for the project from beginning to end” (Morse, Swanson et al. 2001); in this thesis, the research objectives, goal and plan are presented in detail in chapter 3. This research project has been divided into three main phases: Preliminary, Development and Testing. The detailed steps for each phase are presented in chapter 3.

c) Sampling:

“Data collection and analysis proceed until the researcher has collected adequate data – data from different participants, various contexts, and various circumstances and situations – that are similar and fit within the same Category” (Morse, Swanson et al. 2001). Accordingly, the sampling in this thesis has two main purposes:

- For the purposes of developing the top-down evaluation method, a pre-defined set of engineering design classifications has been used to study the design of lightweight SPA methods based on Vincenti’s classifications;
- For the purposes of developing the bottom-up evaluation method, the guidelines to perform systematic literature review in software engineering (Kitchenham 2007) has been used to define: the research questions, search process and recourses, inclusion criteria, exclusion criteria, data to be extracted, data analysis and dissemination activities that led to identifying the sample of references to search for evidences. The process of sampling is iterative. At the beginning of this research project, a set of relevant papers was collected. The search for new relevant publications was continued during the data extraction and categories building phase. Once a new publication is found, all possible evidences are extracted and compared with the categories and evidences already extracted and then add them to the collection of evidences. This process is repeated until no more new evidences are found.

Moreover, as stated in (Morse, Barrett et al. 2002) “By definition, saturating data ensures replication in categories; replication verifies, and ensures comprehension and completeness”. As can be seen in Annex V and Table 7.4, the set of collected evidences in the bottom-up approach includes 38 evidences distributed into 5 main classes. The total frequency for the whole set is 207. Hence, the saturation of the collected data is achieved through the replication in categories already on hand.

d) **Bracketing:**

“Bracketing means that the information learned about prior work is simply put on hold and is not used as a framework or conceptual schema for the proposed study” (Morse, Swanson et al. 2001). In this case, and as mentioned previously, there was no formal evaluation framework based on evaluation theory concepts to evaluate the SPA methods thus forming the impetus for this research project. Moreover, aligning the SPA methods design with the engineering design classifications, based on Vincenti’s classification, has not been discussed previously. Researchers and practitioners experiences in designing and implementing SPA methods have been used to feed the current research with the necessary evidences to build the evaluation method.

e) **Methodology coherence:**

“The aim of methodological coherence is to ensure congruence between the research question and the components of the method. The interdependence of qualitative research demands that the question match the method” (Morse, Barrett et al. 2002). In the research objectives and methodology chapter, the goal for this research project was stated as “Evaluate the success of lightweight software process assessment methods”. To achieve this goal, the question raised is: What are the bases to develop such evaluation methods? Consequently, two different approaches are used to that end: Top-down and Bottom-up approaches. Developing the evaluation methods based on these two approaches is based on the findings from other disciplines outside of software engineering, such as:

- Evaluation theory concepts to provide the framework for the proposed evaluation;

- Engineering design classifications such as Vincenti's classifications.

The development of the evaluation methods is also based on the experiences of researchers and practitioners in the field of SPA to collect the success evidences that affect lightweight SPA methods. The evaluation theory concepts as well as the engineering design classifications have been studied in the specific context of lightweight SPA methods.

Finally, (Morse, Barrett et al. 2002) argued that “the aspect of theory development is to move with deliberation between a micro perspective of the data and a macro conceptual/theoretical understanding”, where theory, in this case, can be developed as an outcome of the research process. Consequently, the outcome of this thesis contributes to the theory building of how one can evaluate the success of lightweight SPA methods.

The outcome of this thesis has been grounded in two main bases:

1. The use of engineering design classifications from an engineering perspective.
2. Collected set of evidences from the literature concerning the successful implementation of SPA methods.

Accordingly, the following section discusses the validation of the outcome of this thesis based on Grounded theory criteria.

8.4 Validation based on grounded theory

Qualitative empirical software engineering research has much in common with social science research such as Cognitive Science and Psychology, (Carver 2007). One of the qualitative research approaches to develop hypotheses iteratively in such fields is the grounded theory method (GTM). Grounded theory was developed by Barney Glaser and Anselm Strauss in the early 60s (Glasser and Strauss 1965) (Glasser and Strauss 1967). As discussed in (Carver 2007), the basic principle behind Grounded theory is that the hypotheses and theories emerge

bottom-up from the data rather than top-down from existing theory. Using this approach, a researcher begins with an existing data set and abstracts a hypothesis or a theory that accurately describes that data. Then, as more data sets become available, the hypotheses and theories are refined to continue to accurately describe all of the extant data.

As proposed by (Glaser 1978) and discussed in (Bryant 2002), the grounded theories “should have fit and relevance, they must work and be readily modifiable”.

By fit, Glaser means that the categories of the theory must fit the data. “Since most of the categories of grounded theory are generated directly from the data, the criterion of fit is automatically met” (Bryant 2002).

By works, Glaser means that the theory can “explain what happened, predict what will happen, and interpret what is happening in the area of substantive or formal inquiry”. A theory will work if its categories fit and if the theory is relevant to the ‘action of the area’ (Bryant 2002).

Glaser stated that a grounded theory automatically achieves relevance because the method allows core ideas to emerge. These ideas are relevant being directly generated from and grounded in the data. The GTM researcher “spends his time modestly, but assertively, searching for and discovering the relevance in his data” (Bryant 2002).

Glaser stated that the modifiability of a grounded theory is embedded in the methods used to create it. Glaser explains that the generation of theory is in fact an ever-modifying process. As new data is collected, the emerging theory is constantly modified and therefore maintains its relevance.

For the proposed evaluation methods which are grounded on the data, the fit criterion is met since the evaluation methods are built based on the collected set of evidences for successful implementation of the SPA method and the engineering design classifications. The Works criterion is also met by the evaluation methods since the methods are based again on the

collected set of evidences for successful implementation of the SPA method and the alignment with the engineering design classification. The hypothesis here is that the more these evidences are met, the more the SPA method is successful. The relevant criterion is met if the fit and work criteria are met. Consequently, the evaluation method is also relevant. The modifiability criterion is also met, since as new data is collected, bringing new requirements or success factors, the evaluation method can be easily changed to include these new evidences.

Here, at the end of this chapter, mention should be made that the face-to-face feedback session held mainly at the end of the first case study presented in the next chapter, as well as the other feedbacks received by email from the other two case studies, also helped in verifying the fitness and relevance of the contents of the proposed evaluation methods.

CHAPTER 9

CONDUCTED CASE STUDIES

9.1 CASE STUDY 1 – EVALUATION OF THE S3^{mAssess} ASSESMENT METHOD

This case study presents an evaluation of an assessment method dedicated to assessing the software maintenance process, namely the S3^{mAssess} Method. Through this case study, the term ‘evaluator’ refers to the developer of the evaluation method proposed in this research, while the expression ‘evaluation participant’ refers to the developer of the assessment method to be evaluated.

9.1.1 Software maintenance process

The software life cycle is divided into two distinct parts (April and Abran 2008): the initial development of software and the maintenance and use of software. Software maintenance is considered as one of the five primary processes in the software life cycle processes of the ISO 12207 international standard. Although software maintenance comprises processes and activities that are not handled by the software development process, software maintenance still uses processes and activities of software development, especially while implementing a modification to existing software applications (April and Abran 2008).

The assessment method used to assess software maintenance processes should assess both the processes specific for the maintenance activities and those processes used in development activities. Moreover, the design process of assessment methods is the same regardless whether the assessment method is to assess development process or the maintenance process. Consequently, the evaluation method devised in this research is also applicable for the evaluation of assessment methods of the software maintenance process.

The assessment and improvement of a software maintenance process is discussed in (April and Abran 2008) whose book documents a software maintenance process model and software maintenance maturity model $S3^m$. Moreover, (April and Abran 2008) propose an assessment method based on the $S3^m$ model. The $S3^m$ model is an outcome of research work documented in (April, Abran et al. 2004; April, Abran et al. 2004; April, Abran et al. 2004; April 2005; April and Desharnais 2005; April, Huffman; et al. 2005; Abran and April 2006; April, Desharnais et al. 2006).

9.1.1.1 Software maintenance maturity model $S3^M$

Organizations can use the $S3^m$ model to launch and sustain a continuous improvement program tailored to software maintainers by initially benchmarking their current maintenance practices against this model. This approach will help maintenance organizations identify their strengths and weaknesses in delivering software maintenance services. With the identification of gaps, maintenance organizations can identify, through a comparison with the model, what issues to address and how to address them, and, by following through, improve their software maintenance processes (April and Abran 2008).

The scope of the $S3^m$ model is to deal with the software maintenance and evolution processes that are under an organization's direct control. A practical approach was used to apply proven knowledge in software maintenance engineering to offer relevant exemplary practices to improve software maintenance (April and Abran 2008). The $S3^m$ model is restricted to small maintenance and evolution activities, and this process model is not appropriate for software maintenance projects of larger scope which should be managed as projects using project management techniques. For the maintenance workload requiring project management expertise and techniques, the CMMi and other maturity models would be more suited. The $S3^m$ model (April and Abran 2008):

1. Is based on the customer's perspective.

2. Is relevant for the maintenance of application software: a) developed and maintained in-house; b) configured and maintained in-house or with a subcontractor's help; and c) outsourced to an outside supplier.
3. Provides references and details for each exemplary practice.
4. Offers an improvement approach based on roadmaps and maintenance categories.
5. Covers international software life cycle processes and maintenance standards like ISO12207, ISO14764, ISO9001, ISO20000 and ISO14764.
6. Covers relevant parts of the CMMi, a reference model for software improvement.

The S3^m includes four process domains, see Figure 9.1:

1. Maintenance process management.
2. Maintenance requests management.
3. Software Evolution engineering.
4. Support to software evolution engineering.

These process domains contain 18 Key Process Areas – KPA and 73 Roadmaps for maintenance practices.

9.1.1.2 The software maintenance assessment method S^{3mAssess}

The S3^m mini-assessment method, named S^{3mAssess}, has been developed to obtain a reliable maturity rating for software maintenance processes without investing too much effort. Each practice in each process specified in the maturity model is used to build the related question for that practice; the whole set of questions is stored in an Excel sheet. Usually, assessment questions are answered through interviews with the software maintenance resource person and senior management.

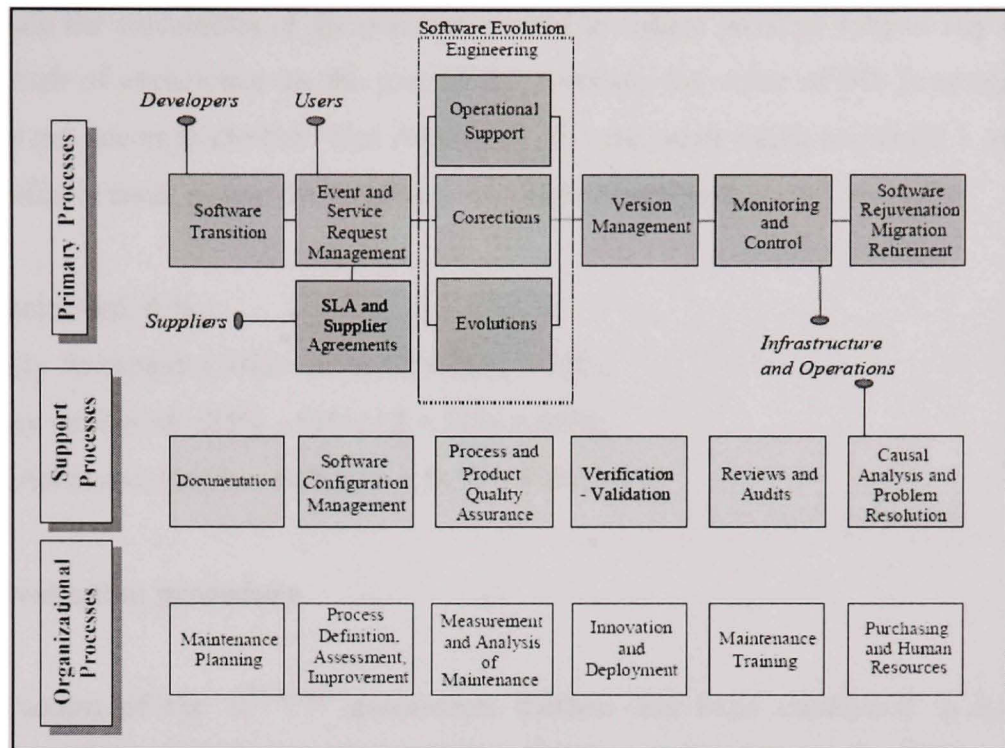


Figure 9.1 S3m process model.
(April, Hayes et al. 2005)

The questions for level 0 are of the form True/False and use the negative form. These questions try to confirm the absence of a specific software maintenance process. For example, “The software maintenance organization does not manage user requests or software events” - Answering “True” to this question leads to a rating of 0% and “False”, to a rating 100%.

For the other maturity levels, the choice of responses is in conformity with ISO 15504:

N: Not Achieved: 0-15%. There is little evidence that the process objectives and goals are met.

P: Partially Achieved: 16%-50%. Some of the objectives and goals are met.

L: Largely Achieved: 51%-85%. Significant portions of the objectives and goals are met.

F: Fully Achieved: 86%-100%. The objectives and goals of the process are fully met.

To facilitate the calculation of the percentages and to reduce possible subjectivity due to the possible lack of experience on the part of the assessor, the value of 0% is assigned if the process is not accomplished or “Not Achieved”. For the other rating levels (P, L and F), the median value is used, giving the following four possible ratings:

N: Not Achieved: 0 %;

P: Partially Achieved: $(50\% - 16\%) / 2 + 16\% = 33\%$;

L: Largely Achieved: $(85\% - 51\%) / 2 + 51\% = 68\%$;

F: Fully Achieved: $(100\% - 86\%) / 2 + 86\% = 93\%$;

9.1.2 Evaluation procedure

The evaluation of the $S^{3mAssess}$ assessment method has been conducted using the two evaluation methods proposed in this research project:

- Bottom-up evaluation method based on success evidences; and
- Top-down evaluation approach based on engineering viewpoint.

The evaluation has been conducted in four phases:

1. Interview (July 12, 2008): The evaluation participant of the $S^{3mAssess}$ assessment method has been interviewed by phone. The evaluation tool has been used to ask questions and record answers.
2. Document review (July 1-25, 2008): different documents and one book related to the maintenance process have been reviewed and includes (April, Hayes et al. 2005; Paquette, April et al. 2006; April and Abran 2008); the $S^{3mAssess}$ version 2006 assessment tool related to the research paper (Paquette, April et al. 2006) has also been reviewed to see how the assessment tool works.
3. Feedback session (Sep. 23, 2008): A detailed feedback session was arranged with the evaluation participant about two months after the initial evaluation to review the content

of the draft of the evaluation report. The evaluation participant's feedback on the evaluation method at the end of the interview was also recorded. A few more references provided by the evaluation participant during the feedback session was also reviewed, as discussed in section 9.1.5.2, in the period (Sep. 25-30, 2008). A final version of the evaluation report was then produced.

4. Submission of evaluation report (Oct. 20, 2008): The interviewee's answers are documented and a draft report showing the strengths and weaknesses was sent to the evaluation participant.

9.1.3 S3^{mAssess} assessment method evaluation results based on the bottom-up evidence-based approach

Based on the evaluation procedure presented in section 1.2, the evaluator conducted an evaluation of the S3^{mAssess} method with the evaluation participant. The summary of the evaluation is shown in Table 9.1. The bottom-up evaluation table consists of the following columns:

- a. Evidence class name.
- b. Total number of evidences in the class.
- c. Strengths points: Calculated as the sum of the evidences satisfied fully, which are assigned weight 1, plus the sum of evidences satisfied partially, which are assigned weight 0.5.
- d. Percentage of strengths: Computed as (Strengths points/total evidence per class)*100%.
- e. Weakness points: calculated as the count of evidences unsatisfied, which are assigned weight 0 plus the sum of evidences satisfied partially, which are assigned weight 0.5.
- f. Percentage of weaknesses: computed as (weakness points/total evidence per class)*100%.
- g. Percentage of strength contribution: This term referred to the total number of evidences that are taken into consideration for each class during the design phase of the method. This percentage is computed as (strengths point per class/total strengths points for all classes)*100%.

As observed from Table 9.1, the $S^{3mAssess}$ method:

- Fulfills most of the success evidences related to the ‘SPA method’ class with about 96% of the evidences for this class;
- While the $S^{3mAssess}$ method did not fulfil most of the evidences related to the ‘user’ class with 80% evidences not fulfilled.

Table 9.1

Bottom-up evaluation results of $S^{3mAssess}$ method

SPA Categories	Total evidences	Strength Points	%	Weakness Points	%	% strength contribution
Method	13	12.5	96.2%	0.5	3.8%	34.7%
Supportive tool	4	3.5	87.5%	0.5	12.5%	9.7%
Procedure	5	3	60.0%	2	40.0%	8.3%
Documentation	9	5.5	61.1%	3.5	38.9%	15.3%
User	5	1	20.0%	4	80.0%	2.8%
Total	36	25.5	70.8%	10.5	29.2%	70.8%

9.1.3.1 Strength points

The $S^{3mAssess}$ method has the following strengths:

1. SPA method:

- Data gathering technique – interview and document review.
- Flexible and customizable in the sense that the method takes all maintenance processes into consideration and the users determine which processes are to be assessed.
- Coverage to a process reference model which is the same reference model as in ISO 15504.
- Identification of strengths, weaknesses, risks and improvement opportunities
- Suggest improvement proposals.
- Publicly available and usable for on-site and self assessment.

- g. Simple and well-structured method having no more than 150 questions in the questionnaire.
- h. Complete method in the sense that the method assesses processes which are supposed to be assessed according to the process model.

2. SPA supportive tool:

- a. The supportive tool covers different assessment phases.
- b. The method generates a semi-automatic assessment report.
- c. Adaptable and flexible based on the fact that the SPA method contains the whole set of possible processes and the tool allows the user to choose which processes to assess.

3. SPA procedure:

- a. Prepare participants in the assessment - through training sessions.
- b. Work to build a relationship of trust and confidence between assessor and sponsors of the assessment.
- c. Produce an assessment report to be delivered it to the organization's representative.
- d. Ensure confidentiality of participants.

4. SPA documentation:

- a. Identification of the assessed organization unit, including the scope of the assessment.
- b. The assessment process is documented.
- c. Documentation of the assessment data and ratings.

9.1.3.2 Weakness points

Several weaknesses exist in the $S3^{mAssess}$ method. These weaknesses correspond to those evidences rated as partially or non adequate in the evaluation and include:

1. SPA method:

- a. Comply with a formal assessment method.

The $S3^{mAssess}$ method is partially compliant with the requirements of ISO 15504. More work is needed to ensure that the $S3^{mAssess}$ is compliant with a formal assessment method.

- b. Ensure reliability.

No studies have been found to discuss the reliability and repeatability of the assessment results. Ensuring reliability will give confidence to the organization to rely on the assessment results and make further decisions for improvements.

2. SPA supportive tool:

Create and use a database of historical assessment data.

The $S3^{mAssess}$ method makes partial use of previous assessment results but it does not build a database of assessment results. Building a database after the assessment process which contains the process profiles and other necessary data would be useful for new assessment trials and also for comparing assessment results with previous assessment results.

3. SPA procedure:

Hold feedback session after an assessment.

After conducting the assessment process, the assessor should hold a feedback session to present the results of the assessments to the organization (that is, the assessor does not limit himself to only delivering the final report to the interviewee). This feedback session also discusses the participants' comments and suggestions about the assessment process. These discussions can enable continuous improvement to the assessment method.

4. SPA documentation:

- a. Documenting the assessment purpose, objectives and resources needed.

The $S3^{mAssess}$ should provide more detailed information about the purpose of the assessment to be conducted, the objectives and resources needed for the assessed organization which is partially supported by the $S3^{mAssess}$ method.

- b. Providing guidance for the assessment team.

The $S3^{mAssess}$ should define the assessment team members' background, experience and their level of knowledge in the assessment method and underlying model. This part of documentation also defines the team leader's management and technical skills related to the assessment process.

- c. Documenting the assessment confidentiality.

The confidentiality procedure that the $S^{3mAssess}$ method is implementing should be documented.

- d. Documenting the templates.

The $S^{3mAssess}$ should document all templates for the necessary documents to be used through the assessment or to be produced at the end of the assessment process to help reduce the effort of reporting assessment results: Confidentiality agreement template, assessment agreement template and final report template.

- e. Providing the guidance for the follow-up meetings.

$S^{3mAssess}$ method should provide guidance for assessors who will conduct the follow-up meetings.

5. SPA user:

- a. Defining the assessment sponsor responsibilities.

$S3^{mAssess}$ should define the responsibilities of the assessment sponsor and interviewees. The assessment sponsor should verify that the assessment will provide the expected results by verifying the skills and experience of the assessment team and by verifying the confidentiality of interviewee and any other related responsibilities.

- b. Defining assessment team responsibilities.

$S3^{mAssess}$ method should document the responsibilities of the assessment team leader (even if the team consists of one person) which include: Ensure that the assessment is conducted in accordance with the underlying assessment method; ensure the commitment of the sponsors; ensure the availability of necessary documents for assessment; ensure that the assessment requirements are met; and ensure the readiness of all participating assessors.

- c. Involving senior management in the assessment process.

$S3^{mAssess}$ should ensure the commitment of senior management to the assessment process and their involvement by attending the meetings, participating in setting priorities for the processes to be assessed and improvement actions.

d. Ensuring the participant commitment.

The $S3^{mAssess}$ method should ensure the commitment of the assessment sponsors to the whole assessment process. The assessment team leader has the responsibility to ensure the sponsor's commitment.

e. Ensuring the sponsor and staff believe the assessment will give results.

To encourage effective involvement in the assessment process, the assessment participants should believe that the assessment will give results and will contribute to the improvement of the organization's behaviour and performance. Hence, the $S3^{mAssess}$ method should also work to build this belief and demonstrate that the assessment method is complete.

For each category, the assessment method should satisfy at least 60%, as a threshold value for success, of the total evidences for that category to be considered as having an acceptable level of success in that category. Hence, as shown in Table 9.1, the 'users' category need more work to achieve the minimum acceptable level of success.

9.1.4 $S^{3mAssess}$ method evaluation results based on Top-down approach

The evaluation participant, has also answered the questions of the top-down evaluation method and referenced the evaluator to the maintenance book (April and Abran 2008) for more details to answer other questions.

As can be observed from Table 9.2, the number of criteria per engineering design class is not equal for all the classes. The participation of each class in the total design effort is not the same.

Based on the evaluation procedure presented in section 9.1.2, the evaluator has conducted the top-down evaluation of the $S^{3mAssess}$ method with the evaluation participant. The summary of the evaluation is shown in Table 9.3.

Table 9.2

Number of questions per engineering design class

Vincenti's design classes	Questions	%
Fundamental design principles	13	46.4%
Criteria and specifications	5	17.9%
Theoretical tools	2	7.1%
Quantitative data	2	7.1%
Practical considerations	3	10.7%
Instrumentalities	3	10.7%
Total	28	100.0%

The collected data in the top-down approach is summarized in Table 9.3, containing the following columns:

- Engineering design class name.
- Total number of criteria in the class.
- Satisfied criteria, calculated as the sum of the criteria satisfied fully, which are assigned weight 1, plus the sum of criteria satisfied partially, which are assigned weight 0.5.
- Percentage of satisfied, computed as (satisfied criteria/total criteria per class)*100%.
- Unsatisfied criteria, calculated as the count of criteria unsatisfied, which are assigned weight 0 plus the sum of criteria satisfied partially, which are assigned weight 0.5.
- Percentage of unsatisfied, computed as (unsatisfied criteria/total criteria per class)*100%.

By analyzing statistics from the evaluation tool, as shown in Table-9.3, one can note the following:

- The $S^{3mAssess}$ method, in its design phase, has covered all the issues mentioned in the “criteria and specification”, “quantitative data” and “practical considerations” of the evaluation tool.
- The $S^{3mAssess}$ method covered about 67% of the issues mentioned in the instrumentalities class.

3. The $S^{3mAssess}$ method covered about 65% of the issues mentioned in the “fundamental design principles” class.
4. The $S^{3mAssess}$ method covered about 50% of the issues mentioned in the “theoretical tools” class.

Table 9.3
Top-down evaluation results of $S^{3mAssess}$ method

Class	Total Criteria	Satisfied	%	Unsatisfied	%
Fundamental design principles	13	8.5	65.4%	4.5	34.6%
Criteria and specifications	5	5	100.0%	0	0.0%
Theoretical tools	2	1	50.0%	1	50.0%
Quantitative data	2	2	100.0%	0	0.0%
Practical considerations	3	3	100.0%	0	0.0%
Instrumentalities	3	2	66.7%	1	33.3%
Total	28	21.5	76.8%	6.5	23.2%

The main issues that the $S^{3mAssess}$ method should address are as follows:

1. The $S^{3mAssess}$ method should conduct an explicit assessment preparation phase during which the following activities should be accomplished
 - a. Identify the organization’s business needs to be used to specify on which processes and practices the assessment method should focus.
 - b. Make use of previous assessment reports (if any).
 - c. Refer to the organizational documents and reports to better understand the organizations needs and current situation.
 - d. The $S^{3mAssess}$ method should produce an assessment plan which clearly defines all steps in the assessment.

- e. Distribute a pre-assessment questionnaire to explore the participants' opinions, needs and expectations from the assessment. The assessor should distribute the questionnaire during a preparation session with the participants and sponsors of the organization to discuss the assessment process and the results from the activities above.
 - f. Based on the final results of the preparation sessions and pre-assessment questionnaire, an assessment initiation file could be developed that includes sponsor commitment and definition of the input data such as business needs, reports from previous assessments, organizational documents, assessment tools and any other necessary information for launching the assessment.
 - g. More tracking for the evidences of process performance and capability should be done by the $S^{3mAssess}$ method.
2. Define and document explicitly the different sub-divisions that constitute the $S^{3mAssess}$ method (for example: prepare for the assessment, conduct assessment, analyze results, reporting) thereby dividing the design phase into several phases based on these sub-divisions.
 3. A specific procedure should be defined to improve the $S^{3mAssess}$ method (for example: applying an iterative approach to improve the method).

The strengths and weaknesses explored by the two methods (top-down and bottom-up) are not necessarily distinct and may overlap. To get the best results, the recommendations generated by the two methods should both be taken into consideration while building a new SPA method or upgrading existing methods.

Since the $S^{3mAssess}$ method is already designed, the set of recommendations provided in this evaluation should be taken into consideration in the next version of this assessment method.

9.1.5 Results of the feedback session

As mentioned previously, a quick feedback has been recorded at the end of the interview with the participant and a detailed feedback session was arranged two months later conducting an evaluation to review the validity of the content of the 'Draft Evaluation Report'. The results of these two feedbacks are discussed in the following paragraphs.

9.1.5.1 Quick feedback results

At the end of the phone evaluation, the interviewee expressed his opinion on reviewing the evaluation method; the interviewee's main comment was that an evaluation should typically be conducted with a third party, such as the users of the assessment method rather than with the designers of the assessment methods to ensure unbiased answers.

The answers to this concern are as follows:

1. The users of the evaluation methods proposed in this research project are mainly the designers of the SPA methods (and any one interested in measuring the success of an SPA method).
2. The proposed evaluation methods contain some technical terms and concepts which may not be known by an ordinary user of any SPA method who has been trained to use the assessment method but who has not been trained in the design - for example, the SPA method is based on which process reference model and which process assessment model? Moreover, determining whether what is provided by the current version of the SPA method is enough or whether more work should be done; and whether this work is fully, partially or not adequate. This information may not be known by the users of the SPA method.
3. When conducting software process assessment in an organization, the assessor usually reviews documentation and asks people in the organization about their software process to arrive at the assessment results (strengths, weaknesses and recommendations). The same is done when evaluating SPA methods. The publications and documents

explaining the SPA method in question are reviewed and the designers are interviewed to discuss their SPA methods.

9.1.5.2 Detailed feedback results

The Draft report was sent to the evaluation participant and a feedback session was conducted two months after the evaluation. During the feedback session, the evaluation report findings were presented and strengths and weaknesses were discussed. The participants of this feedback session were:

- Mr. Alain April in the role of the evaluation participant who have conducted the evaluation during the interview (interviewee);
- Mr. Mohammad Zarour in the role of the evaluator of the $S^{3mAssess}$ method and the one who conducted the interview as the interviewer;
- Mr. Alain Abran as an audience and observer of the feedback session.

While discussing the results of conducting the bottom-up evaluation method, the main points raised during this discussion can be summarized in the following paragraphs.

1. The evaluation participants agreed that the strengths points have been adequately considered in the $S^{3mAssess}$ method.
2. For the weakness points, the participants mentioned that an updated version of the assessment method had been developed with inputs from a research work done by a Master's student in Belgium (Vincent 2008) which aims at developing an assessment method for the S^{3m} model that complies with the ISO15504 standard. The proposed assessment method in (Vincent 2008) has been reviewed by the evaluator (Mohammad Zarour) to explore weakness points that were addressed in the updated version of $S^{3mAssess}$ method as in (Vincent 2008).

The research conducted in Belgium to develop a new assessment method was done independently of this research work, using a different methodology and from a different viewpoint (that is, from the users' viewpoint of that method, rather than from the

designer's viewpoint (or an expert's viewpoint). The updated assessment method has taken into consideration some of the weakness points mentioned in this project's evaluation report and include:

- a. Comply with a formal assessment method: the participant mentioned that the work done by (Vincent 2008) defined a new assessment method which complies with the ISO 15504 standard.
 - b. Build a database of historical data: the new version of the evaluation tool has developed a database of historical data, but, as mentioned by the evaluation participants, it is not in use yet. Hence, this weakness point still holds since the aim of building the database is to use the historical data.
 - c. Feedback session: the new S3^m assessment method proposes holding feedback sessions known as "Brainstorming".
 - d. Documentation: the new S3^m assessment method proposed in (Vincent 2008) has documented the following:
 - The assessment purpose, objectives and resources needed;
 - Provide guidance for the assessment team's backgrounds, skills and responsibilities.
 - e. Define the assessment team responsibilities.
3. The evaluation participant was puzzled about the appearance of contradictions to the following two assessment results in the Draft report:
- a. Strength = 'ensuring the accuracy of findings' in section 1.3.1.1 of the draft report.
 - b. Weakness = 'ensuring the reliability of assessment results' in section 1.3.1.2 of the draft report.

The references that mentioned the reliability of results as a requirement for the SPA method for SMEs are mainly (Anacleto, Wangenheim et al. 2004), (Laporte, Desharnais et al. 2005) and (Wiegers and Sturzenberger 2000), while the reference that talked about the accuracy and validation of findings was (CMMI-SUT 2006) and (ISO/IEC 2003).

- Accuracy: The accuracy of findings is concerned with providing a mechanism to consolidate the collected data into accurate findings based on certain criteria, (ensure that findings are derived from objective evidences seen or heard during

data collection and the data are worded without attribution). Usually this consolidation is achieved by presenting the collected data results to the organization's participants;

- **Reliability:** Reliability is concerned with ensuring that the produced assessment results at the end of the assessment (not the data collected) are consistent and repeatable. This ambiguity can be removed by providing more details about each result; hence, the description of each of these two evidences is enhanced to reflect this difference.
4. The participant was asked to provide examples of the possible templates that should be provided while conducting the assessment method. As a result, the description of this weakness point has been modified to include examples of the possible templates, including a confidentiality agreement template, assessment agreement template, and final report template.

Other weakness points that are mentioned in the bottom-up evaluation method and not taken into consideration in the updated assessment method proposed in (Vincent 2008) are still considered as weaknesses that need further study.

As a result, the modified evaluation results updated with the findings of the feedback session can be seen in Table 9.4.

Table 9.4

Modified Bottom-up evaluation results of S^{3mAssess} method - based on feedback session

SPA Categories	Total evidences	Strengths Points	%	Weakness Points	%	% strength contribution
Method	13	12.5	0.5	96.2%	3.8%	34.7%
Supportive tool	4	3.5	0.5	87.5%	12.5%	9.7%
Procedure	5	4	1	80.0%	20.0%	11.1%
Documentation	9	6.5	2.5	72.2%	27.8%	18.1%
User	5	2	3	40.0%	60.0%	5.6%
Total	36	28.5	7.5	79.2%	20.8%	79.2%

Regarding the weakness points identified through the top-down evaluation method, the discussion can be summarized as follows:

1. The evaluation participant mentioned that the updated assessment method of (Vincent 2008) has also taken into consideration the following points:
 - a. Conducting an explicit assessment preparation phase: the second phase of the mini-assessment method is a preparation of the assessment step; the details of this step is given in (Vincent 2008).
 - b. Refer to the organizational documents and reports to better understand the organization's needs and current situation.
 - c. Producing an assessment initiation file which records the assessment input.
 - d. Produce an assessment plan.
 - e. Define and document the sub-divisions that constitute the $S^{3mAssess}$ method.
2. The weakness point in "tracking for the evidences of process performance and capability": The evaluation participant mentioned that this point is not supported by the assessment method because the underlying assessment model lacks process capability elements to accomplish this task. Hence, it is difficult to support this point.

Other weakness points that are mentioned in the top-down evaluation method and not taken into consideration in the updated assessment method proposed in (Vincent 2008) are still held as weaknesses that need further study.

In summary, Vincent in (Vincent 2008) has built an updated version of the $S^{3mAssess}$ assessment method based on the $S3^m$ model as presented in (April and Abran 2008). The new updated method complies with the ISO 15504. The weaknesses of the old version which have been addressed by Vincent are moved from the weaknesses side to the strengths side of the updated method, while the weaknesses in the old version, which have not been addressed by Vincent, are propagated to the updated method. Table 9.5 summarizes the results of the modified evaluation based on feedback findings as discussed above.

Based on the discussions in the feedback session, the bottom-up evaluation method has been changed slightly as follows:

1. In the documentation class, the two evidences “provide the necessary documents” and “document necessary templates” are both merged into single evidence as “Provide all necessary documents and templates”.
2. To comply with the evidences of the method class, two evidences related to the “tool support” class have been added and are already presented in section 7.3.3.2 as:
 - a. The tool supports the confidentiality of the assessment.
 - b. The tool ensures the repeatability of results.
3. New evidence in the “users” class has been added as “Credibility: the assessment team believes the assessment will give results”. The description of this evidence is presented in section 7.3.3.5.

In summary, the evaluation participant agreed:

1. With all the strengths identified.
2. Concerning all the weaknesses identified at the time of the evaluation.
3. And confirmed that some of these weaknesses were significant enough that the participant himself had implemented remedial actions since the interview occurred (thereby confirming the validity of these weaknesses at the time of the evaluation).

Table 9.5

Modified Top-down evaluation of S3^{mAssess} method – based on feedback session

Class	Total Criteria	Satisfied	%	Unsatisfied	%
Fundamental design principles	13	11.5	88.5%	1.5	11.5%
Criteria and specifications	5	5	100.0%	0	0.0%
Theoretical tools	2	1	50.0%	1	50.0%
Quantitative data	2	2	100.0%	0	0.0%
Practical considerations	3	3	100.0%	0	0.0%
Instrumentalities	3	2	66.7%	1	33.3%
Total	28	24.5	87.5%	3.5	12.5%

9.2 CASE STUDY 2 – EVALUATION OF THE ‘MICRO-EVALUATION’ METHOD

This case study presents an evaluation of Micro-Evaluation assessment method.

9.2.1 Introduction to the ‘Micro-Evaluation’ method

The OWPL model has been designed with respect to the particular context of small businesses to help them to improve their software practices accordingly. The structure of the OWPL model involves processes, practices and success factors and defines 10 processes: requirements management, project planning, project tracking and oversight, development, documentation, testing, configuration management, sub-contractors management, quality management, and experience capitalisation process. Each of these processes has a number of practices. The OWPL gradual approach is based on a three-stage software process improvement framework. The three stages are a) Micro-evaluation assessment b) OWPL assessment and c) SPICE or CMMI assessment.

At the first stage, a very simplified questionnaire called the ‘micro-evaluation’ is used to collect information about the current software practices in small organizations and to make people sensitive to the importance of software quality aspects. This questionnaire covers six key axes selected on the basis of former experience with SME evaluation as the most pertinent to the targeted organisations. These axes are: Quality assurance, Customer management, Subcontractor management, Project management, Product management, and Training and Human Resources management.

The Micro-evaluation involves one person in the evaluated organisation. This person must either have sufficient knowledge of software quality matters or already be in charge of software quality. The information collected is then used as a starting point to determine the goals of a more accurate evaluation according to the OWPL model which has been developed

on the same bases. The organizations showing a certain maturity level are encouraged to conduct SPICE or CMM assessment.

9.2.2 Evaluation procedure

The evaluation of the Micro-evaluation assessment method has been conducted using two evaluation methods:

- Bottom-up evaluation method based on success evidences;
- Top-down evaluation approach based on engineering viewpoint.

The evaluation has been conducted in three phases:

1. Interview (Aug. 21, 2008): Mr. Alexandre Simon (the ‘interviewee’), One of the authors and main researchers in the design and implementation of the Micro-Evaluation assessment method, was contacted and asked to answer the evaluation questions provided in the evaluation tool. The evaluation tool was forwarded by email, and the interview was conducted by phone by Mr. Desharnais.
2. Document review (Aug. 1–31, 2008): different articles which present and discuss the OWPL and micro-evaluation method have been reviewed, including (Habra, Eustache Niyitugabira et al. 1999; Habra, Niyitugabira et al. 1999; Renault 1999; Habra, Renault et al. 2002; Habra and Renault 2004; Laporte, Desharnais et al. 2005; Alexandre, Renault et al. 2006; FUNDP-CETIC 2006; Stambollian, Habra et al. 2006).
3. Submit evaluation results (Oct. 31, 2008): The final report was sent to the interviewee who was asked to return any feedback and comments.

9.2.3 Evaluation of the Micro-Evaluation method based on the Bottom-up approach

Based on the evaluation procedure presented in section 9.2.2, Mr. Desharnais contacted Mr. Simon to conduct the evaluation of Micro-Evaluation method. The summary of the evaluation is shown in Table 9.6.

The evaluation results have been translated into percentages. For each question, a weight has been assigned: 1 for a fully adequate rating, 0.5 for a partially adequate and 0 for a non-adequate. The strengths points are calculated by summing up the number of fully adequate evidences and the sum of the partially adequate evidences. Similarly, the weakness points are calculated by summing up the number of non-adequate evidences and the sum of partially adequate evidences. The percentages are then calculated for the strength points and weakness points.

Table 9.6

Bottom-up evaluation method results of the Micro-Evaluation method

SPA Categories	Total evidences	Strengths Points	%	Weakness Points	%	% strength contribution
Method	13	7	53.8%	6	46.2%	19.4%
Supp. tool	4	2	50.0%	2	50.0%	5.6%
Procedure	5	3.5	70.0%	1.5	30.0%	9.7%
Documentation	9	6.5	72.2%	2.5	27.8%	18.1%
User	5	3	60.0%	2	40.0%	8.3%
Total	36	22	61.1%	14	38.9%	61.1%

As observed from Table 9.6, the Micro-Evaluation method fulfils mostly the success evidences related to the ‘SPA documentation’ category with about 72% of the evidences for this category, while the Micro-Evaluation method fulfilled only half of the evidences related to the ‘supportive tool’.

9.2.3.1 Strength points

The evaluation results show that the Micro-evaluation method has several strength points. These strengths can be summarised as:

1. SPA method:

- a. Data gathering technique – interview: the Micro-evaluation collects the data through a scheduled interview with the organization participants.
- b. Flexible and customizable method that allows adding new processes to be assessed based on the organization's needs.
- c. Identification of strengths, weaknesses, risks and improvement opportunities.
- d. Suggest improvement action plan to start an improvement process.
- e. Simple and well-structured method having no more than 150 questions in the questionnaire.

2. SPA supportive tool:

- a. The supportive tool is useable for the assessor and covers different assessment phases.
- b. Adaptable and flexible tool that easily allows adding new processes for the assessment process.

3. SPA procedure:

- a. The method works to build a trust and confidence relationship between assessors and organization participants.
- b. Producing an assessment report to deliver to the organization representative.
- c. The assessment procedure ensures the confidentiality of the participants.

4. SPA documentation:

- a. The assessment purpose, objectives and needed resources are all documented.
- b. Identification of the assessed organization unit.
- c. The confidentiality of the assessment is documented.
- d. The method documents all necessary documents and templates.
- e. Documentation of the assessment data and ratings.

5. SPA users:

- a. The responsibilities of the assessment participants are defined.
- b. Senior management is involved in the assessment process.

9.2.3.2 Weakness points

The evaluation criteria that are not fully met in the assessment method are considered as weakness points. Although in the Micro-evaluation method and its underlying model OWPL “emphasis more in training than on documentation and formalization aspects” (Habra, Eustache Niyitugabira et al. 1999; Habra, Niyitugabira et al. 1999; Renault 1999; Habra, Renault et al. 2002; Habra and Renault 2004; Laporte, Desharnais et al. 2005; Alexandre, Renault et al. 2006; FUNDP-CETIC 2006; Stambollian, Habra et al. 2006), the un-emphasised aspects should still be presented in a minimum acceptable level.

Hence, and based on the evaluation results, the following points are found to be weakness points and need to be handled by the designers of the Micro-evaluation method.

1. SPA method:

- a. Studying the accuracy of findings.

The collected data should be consolidated into accurate findings according to defined criteria. For the Micro-evaluation method no studies have been found to discuss the consolidation of findings into accuracy of results.

- b. Coverage to process reference model.

The selection of processes that the Micro-Evaluation can assess should be based on a process reference model.

- c. Publicly available and usable assessment method.

- d. Comply with formal assessment method.

Micro-evaluation method, as a lightweight SPA method, should comply with a comprehensive assessment method.

- e. Ensure the reliability of the assessment results.

Although the interview mentioned that the reliability is ensured through numerous case studies, no formal studies have been found to measure to what degree the assessment results produced by these case studies are repeatable. Ensuring repeatability will give confidence to the organization for relying on the assessment results and make further decisions for improvements.

- f. Ensure completeness.

No study was found to discuss the completeness of the Micro-Evaluation method showing that the assessment method has taken into account the essential elements to assess each process in the assessment scope and give all needed results.

2. SPA supportive tool:

- a. Create and use a database of historical assessment data.

No evidence was found which shows that the Micro-Evaluation method builds or uses previous assessment results. Building a database after the assessment process which contains the process profiles and other necessary data would be useful for new assessment trials and also for comparing assessment results with previous assessment trial results.

- b. Generating a semi-automatic assessment report.

The final report of the Micro-Evaluation only makes use of the graphs produced by the assessment tool. Providing as assessment supportive tool which automates more parts of the produced assessment report will produce a more efficient assessment process.

3. SPA procedure:

- a. Preparing participants in the assessment.

The Micro-Evaluation should conduct a preparation phase for the assessment process which includes preparing the participants in the assessment process, including assessors and sponsors to conduct and to participate in the assessment. Preparation also includes the development of an assessment plan specifying the assessment activities, schedules, necessary resources and other issues that may

affect the assessment process. The aim of this preparation is to create a sufficient consensus of the proposed assessment including:

- Clarifying the purpose, scope and method of the assessment;
- The roles and responsibilities of the participants;
- Confidentiality of conducted assessments;
- Proposed schedules and activities of the assessment.

b. Holding a feedback session.

The interview mentioned that feedback is achieved through the last two questions in the questionnaire. A feedback session is usually conducted after the assessment process, not as part of it. Hence, the assessor should arrange a feedback session after the assessment process to present the assessment results and discuss these results and the whole assessment process to get the participants' feedback directly.

4. SPA documentation:

a. Providing guidance for the assessment team.

The Micro-Evaluation should define the assessment team members' backgrounds, experience and their level of knowledge in the assessment method and underlying model. This part of documentation also defines the team leader's management and technical skills related to the assessment process.

b. Documenting the assessment process.

The Micro-Evaluation should document the whole assessment process:

- Definition of the assessment method and underlying model;
- Description of the assessment process;
- Guidance for process selection.

c. Providing the guidance for the follow-up meetings.

The Micro-Evaluation method should provide guidance for assessors who will conduct the follow-up meetings.

5. SPA user:

a. Defining assessment team responsibilities.

The Micro-Evaluation method should document the responsibilities of the assessment team leader (even if the team consists of one person) which include:

Ensure that the assessment is conducted in accordance with the underlying assessment method, ensure the commitment of the sponsors; ensure the availability of necessary documents for assessment; ensure that the assessment requirements are met; and ensure the readiness of any participating assessors.

b. Ensuring participant commitment.

The Micro-Evaluation method should ensure the commitment of the assessment sponsors to the whole assessment process; the assessment team leader is responsible to ensure the sponsor's commitment.

c. Ensuring the sponsor and staff believe the assessment will give results.

Another important aspect that would encourage effective involvement in the assessment process is the belief that the assessment will give results and will participate in improving the organization's behaviour and performance.

By collecting statistics from the evaluation tool, as shown in Table 9.6, one notes the following:

1. The strength points related to the 'method' class gained the most contribution of the total strengths of the Micro-evaluation method with a total contribution of about 19%, having around 54% of the evidences in this class being achieved.
2. The strength points related to the 'supportive tool' class gained the minimum contribution of total strengths of the Micro-evaluation method with a total contribution of about 6%, having 50% of the evidences in this class being achieved.
3. The strength points related to the 'documentation' class gained about 18% of the total strength contribution, achieving about 72% of the evidences in this class.
4. The strength points related to the 'procedure' class gained about 10% of the total strength contribution, achieving about 70% of the evidences in this class.
5. The strength points related to the 'user' class gained about 8% of the total strength contribution, achieving 60% of the evidences in this class.

For each category, the assessment method should satisfy at least 60% of the total evidences for that category to be considered an acceptable level of success in that category. Hence, as shown by the data in Table 9.6, the ‘methods’ and ‘supportive tools’ categories need more work to achieve the minimum acceptable level of satisfaction.

9.2.4 Micro-Evaluation assessment method evaluation results based on Top-Down evaluation method

The evaluation has also been conducted based on the Top-down evaluation method. The questions in this second evaluation method, which is based on Vincenti’s view of engineering design, determine the different decision points that should be answered during the design phase of the assessment method. Since the Micro-evaluation method is already designed, the set of recommendations provided here, as well as other recommendations provided by the first evaluation method, should be taken into consideration in the next version of this assessment method. These recommendations include:

1. The Micro-Evaluation method should be based on a process reference model which defines the processes and their corresponding practices.
2. The Micro-Evaluation method should be based on a comprehensive assessment model that defines ratings, scales and measures for the assessment process.
3. The Micro-evaluation method should conduct an explicit assessment preparation phase during which the following activities should be accomplished:
 - a. Refer to the organizational documents and reports to better understand the organization’s needs and current situation.
 - b. Conduct a preparation session with the participants and sponsors from the organization to discuss the results and findings from the previous steps. A pre-assessment questionnaire may be used to explore the participants’ opinions, needs and expectations from the assessment.
 - c. Based on the final results of the preparation sessions and pre-assessment questionnaire, an assessment initiation file should be developed that includes the

sponsor commitment and definition of the input data such as business needs, reports from previous assessments, organizational documents, assessment tools and any other necessary information for launching the assessment.

- d. An assessment plan should be produced at the end of the assessment preparation phase that describes how the assessment process will proceed.
4. At the end of the assessment, an assessor record should be maintained. The assessor record contains detailed information about the person who conducted the assessment such as: name, credentials and responsibilities.
5. Assessment method should track the organization's evidences of process performance and capability.
6. The assessment results should be formalized as improvement proposals or an action plan to start the improvement process.
7. Define and document explicitly the different sub-divisions that constitute the micro-evaluation method dividing the design phase into several phases based on these sub-divisions.
8. Through the study of the different publications and reports discussing the Micro-evaluation method to specify the rating mechanism used by this method, the rating mechanism for both practices and processes were not found to be documented clearly in the references provided. The rating mechanism should be defined and documented clearly to help new assessors understand this method.
9. A specific procedure should be defined to improve the micro-evaluation method such as applying an iterative approach to improve the method.

By collecting statistics from the evaluation tool, as shown in Table 9.7, the following is noted:

1. Micro-Evaluation method designers have covered all the issues mentioned in the 'criteria and specification' and 'quantitative data'.
2. The micro-evaluation method covered about 17% of the issues mentioned in the 'instrumentalities' class.

3. The micro-evaluation method covered about 75% of the issues mentioned in the ‘theoretical tools’ class.
4. The micro-evaluation method covered about 50% of the issues mentioned in the ‘practical considerations’ class.
5. The micro-evaluation method covered about 39% of the issues mentioned in the ‘fundamental design principles’ class.

Table 9.7

Top-down evaluation method results of Micro-Evaluation method

Class	Total Criteria	Satisfied	%	Unsatisfied	%
Fundamental design principles	13	5	38.5%	8	61.5%
Criteria and specifications	5	5	100.0%	0	0.0%
Theoretical tools	2	1.5	75.0%	0.5	25.0%
Quantitative data	2	2	100.0%	0	0.0%
Practical considerations	3	1.5	50.0%	1.5	50.0%
Instrumentalities	3	0.5	16.7%	3	83.3%
Total	28	15.5	55.4%	12.5	44.6%

Clearly, the ‘fundamental design principles’ as well as ‘practical considerations’ and ‘instrumentalities’ need more attention and improvements.

9.3 CASE STUDY 3 – EVALUATION OF THE MARES ASSESSMENT METHOD

This case study presents an evaluation of MARES assessment method.

9.3.1 Introduction

The MARES (A Methodology for Software Process Assessment in Small Software Companies) model has been built by researchers from UNIVALI University and the CenPRA research center in Brazil. MARES is designed to support process improvement in the context of small software organizations considering their specific characteristics and limitations, this MARES model is built in conformity to ISO 15504 (ISO/IEC 2003-2006). MARES enhanced the process assessment model mainly by integrating a context-process model in order to support the selection of relevant processes and a process-risk model to support the identification of potential risks and improvement suggestions. The MARES assessment method is divided into five main parts:

1. Planning:

In this phase the assessment is organized and planned. At the end of this phase, the resulting assessment plan is revised and documented.

2. Contextualization:

In this phase the organization is characterized in order to understand its goals, products and its software process. Questionnaires and interviews are used as a means to collect data.

3. Execution:

The selected processes are assessed in detail.

4. Monitoring and control:

All activities during the assessment are monitored and controlled. Corrective actions are initiated, if necessary, and the plan is updated accordingly.

5. Post-mortem:

Once the assessment is finished, a brief post-mortem session is held among the assessors to discuss the performance of the assessment.

9.3.2 Evaluation procedure

Again, the evaluation of the MARES assessment method has been conducted using two evaluation methods:

- Bottom-up evaluation method based on success evidences;
- Top-down evaluation approach based on engineering viewpoint.

The evaluation has been conducted in three phases:

1. Answering the Questionnaire (May 5, 2008 – Aug. 14, 2008): Ms. Wangenheim, one of the authors and main researchers in the design and implementation of the MARES assessment method, was contacted the first time on May 5, 2008 by email and asked to answer the evaluation questions provided in the evaluation tool. Due to her engagement in other activities and conferences the evaluation result was received from her on Aug. 14, 2008.
2. Document review (Aug. 1-31, 2008): different articles which present and discuss the MARES assessment method have been reviewed and include (Anacleto, Wangenheim et al. 2004; Anacleto, Wangenheim et al. 2004; Wangenheim, Anacleto et al. 2004; Wangenheim, Varkoi et al. 2005; Wangenheim, Varkoi et al. 2006; Wangenheim, Weber et al. 2006; Wangenheim, Anacleto et al. 2006).
3. Evaluation report submission (31 Oct, 2008): The final report was sent to Ms. Wangenheim who was asked to send back her comments and suggestions, if any.

9.3.3 MARES assessment method evaluation results based on Bottom-up evaluation method

Based on the evaluation procedure presented in section 9.3.2, Ms. Wangenheim was contacted by email and asked to conduct the evaluation of the MARES method. The summary of the evaluation is shown in Table 9.8.

The evaluation results have been translated into percentages. For each question, a weight has been assigned: 1 for fully adequate rating, 0.5 for partially adequate and 0 for non-adequate. The strengths points are achieved by summing up the number of fully adequate evidences and the sum of the partially adequate evidences. Similarly, the weakness points are achieved by summing up the number of non-adequate evidences and the sum of the partially adequate evidences. The percentages are then calculated for the strength points and weakness points.

Table 9.8

Bottom-up evaluation method results of MARES assessment method

SPA Categories	Total evidences	Strength Points	%	Weakness Points	%	% strength contribution
Method	13	10	76.9%	3	23.1%	27.8%
Supp. tool	4	0	0.0%	4	100.0%	0.0%
Procedure	5	5	100.0%	0	0.0%	13.9%
Documentation	9	7.5	83.3%	1.5	16.7%	20.8%
User	5	4.5	90.0%	0.5	10.0%	12.5%
Total	36	27	75.0%	9	25.0%	75.0%

As observed from Table 9.8, the MARES method fulfilled totally the success evidence related to the ‘procedure’ class and fulfils mostly the success evidences related to the ‘users’ class with 90% of the evidences for this class, while the MARES method did not fulfill any of the evidences related to the ‘supportive tool’ since such a tool is not yet provided for the MARES method.

9.3.3.1 Strength points

The evaluation results show that the MARES method has several strength points; these strengths can be summarised as:

1. SPA method:

- a. Data gathering technique – interview: collects the data through scheduled interviews with the organization participants.
- b. Flexible and customizable method that allows adding new processes to be assessed based on the organization's needs.
- c. Coverage to a process reference model. The MARES process dimension has been developed based on ISO 15504-5 processes which are based on the ISO 12207. But, due to the characteristics of small organizations, some processes have been disregarded as being irrelevant in most cases. If any of these disregarded processes turn out to be important, they are re-integrated based on the ISO 15504-5 as discussed in (Wangenheim, Anacleto et al. 2004).
- d. Identification of Strengths, weaknesses, risks and improvement opportunities.
- e. Suggest improvement action plan to start an improvement process.
- f. The method is publicly available including the process description and artifacts templates. The method can be used in on-site/self assessment, yet requires an experienced assessor, who must be available on-site.
- g. Comply with a comprehensive assessment method: MARES is compliant with the assessment requirements as stated in ISO 15504. The MARES method assesses subsets of the processes that are relevant to the organization's needs.
- h. Simple and well-structured method having no more than 150 questions in the questionnaire.

2. SPA procedure:

- a. Prepare the participant in the assessment. The MARES method, ISO 15504, the developed assessment plan and schedule are presented briefly to all assessment participants at the beginning of the assessment.

- b. The method works to build trust and confidence relationships between assessors and organization participants through the assessment briefing in the beginning as well as a formal term of confidentiality signed by all assessors and the sponsor(s).
- c. Produce an assessment report to be delivered to the organization representative.
- d. The assessment procedure ensures the confidentiality of the participants by signing a confidentiality agreement. The confidentiality of any data provided is guaranteed to all participants of the assessment.
- e. Hold feedback sessions after the assessment feedback is provided through a satisfaction questionnaire to be filled out by the sponsor. A post-mortem meeting is held between the assessors to discuss the performance of the assessment. Although this success evidence is evaluated to be fully adequate in the MARES method, at the end of the assessment, a feedback session with the assessment participants like sponsors and organization representatives should be held to discuss not only the assessment results, but also to discuss, face to face, the assessment method and all issues addressed in the satisfaction questionnaire to get feedback directly, one should not rely only on the satisfaction questionnaire only.

3. SPA documentation:

- a. The assessment purpose, objectives and needed resources are all documented.
- b. Identification of the assessed organization unit.
- c. The confidentiality of the assessment is documented.
- d. The method documents all necessary documents and templates.
- e. Document the assessment process as a whole.
- f. Document of the assessment data and ratings.

4. SPA users:

- a. The responsibilities of the assessment participants are defined.
- b. The responsibilities of the assessment team members are defined.
- c. Senior management is involved in the assessment process.
- d. The participant commitment is ensured through the assessment briefing in the preparation phase.

9.3.3.2 Weakness points

The evaluation criteria that are not fully met in the assessment method are considered as weakness points. Therefore, based on the evaluation results, the following points are found to be weakness points and needs to be handled by the designers of the MARES method.

1. SPA method:

- a. Studying the accuracy of findings.

The collected data should be consolidated into accurate findings according to defined criteria. For the MARES method, no study has been found to define and test such criteria.

- b. Ensure the reliability of the assessment results.

Although the interview mentioned that reliability is ensured through numerous case studies, no study has been found to measure to what degree the assessment results produced by these case studies are repeatable. Ensuring repeatability will give confidence to the organization to rely on the assessment results and make further decisions for improvements.

- c. Ensure completeness.

No study was found to discuss the completeness of the MARES method showing that the assessment method has taken into account the essential elements to assess each process in the assessment scope and give all needed results.

2. SPA supportive tool:

So far there is no tool support for the MARES method. A supportive tool that provides the following features is needed to achieve a successful assessment method:

- a. A tool that is usable and cover the different phases of the assessment.
- b. Create and use a database of historical assessment data.

Building a database after the assessment process which contains the process profiles and other necessary data would be useful for new assessment trials and also for comparing assessment results with previous assessment trials results.

- c. Generating a semi-automatic assessment report.

A supportive tool which produces parts of the final assessment report well produces a more efficient assessment process.

- d. Adaptable and flexible assessment tool.

The supportive tool should allow adding or removing new processes to be assessed to fit the needs and goals of the assessed organization.

3. SPA documentation:

- a. Providing guidance for the assessment team.

MARES should define the assessment team members' backgrounds, experience and their levels of knowledge in the assessment method and underlying model. This part of documentation also defines the team leader's management and technical skills related to the assessment process.

- b. Provide the guidance for follow-up meetings.

The MARES method should provide guidance for assessors who will conduct the follow-up meetings after conducting the improvement phase.

4. SPA user:

Ensuring the sponsor and staff believe the assessment will give results.

Another important aspect that would encourage effective involvement in the assessment process is the belief that the assessment will give results and will lead to the improvement of the organization's behaviour and performance. The MARES method should work to build this belief before and during the assessment.

By collecting statistics from the evaluation tool, as shown in Table 9.8, the following can be noted:

1. The strength points related to the 'method' class gained the most contribution of the total strengths of the MARES method with a total contribution of about 28%, having about 77% of the evidences in this class achieved.

2. The strength points related to the 'supportive tool' class gained the minimum contribution of the total strengths of the MARES method with a total contribution of 0%.
3. The strength points related to the 'documentation' class gained about 21% of the total strength contribution, achieving about 83% of the evidences for this class.
4. The strength points related to the 'procedure' class gained about 14% of the total strength contribution, achieving 100% of the evidences for this class.
5. The strengths points related to the 'user' class gained about 13% of the total strength contribution, achieving 90% of the evidences for this class.

For each category, the assessment method should satisfy at least 60% of the total evidences for that class to be said that it has an acceptable level of success in that category. As shown by the data in Table 9.8, the 'supportive tool' class needs more work to achieve the minimum acceptable level of satisfaction.

9.3.4 MARES assessment method evaluation results based on Top-Down evaluation method

The evaluation has also been conducted based on the Top-down evaluation method. Since the MARES method is already designed, the set of recommendations provided here, as well as other recommendations provided by the first evaluation method, should be taken into consideration in the next version of this assessment method. These recommendations include:

1. The MARES method assessment during the preparation and planning phase should consider the following issues:
 Make use of previous assessment reports: previous assessment reports would give a clear view of the previous states of the organization, the weaknesses previously diagnosed as well as improvement proposals suggested for the organization, all of which would help better understand the organization's current needs and situation.

2. At the end of the assessment, an assessor record should document the assessor name, credentials and related information.
3. The assessment method should track the organization's evidences of process performance and capability ratings.
4. The assessment results should be formalized as improvement proposals or an action plan to start the improvement process.
5. A specific procedure should be defined to improve the MARES method, for example, applying an iterative approach to improve the method.

Table 9.9

Top-down evaluation method results of MARES method

Class	Total Criteria	Satisfied	%	Unsatisfied	%
Fundamental design principles	13	10	76.9%	3	23.1%
Criteria and specifications	5	5	100.0%	0	0.0%
Theoretical tools	2	2	100.0%	0	0.0%
Quantitative data	2	2	100.0%	0	0.0%
Practical considerations	3	2	66.7%	1	33.3%
Instrumentalities	3	2	66.7%	1	33.3%
Total	28	23	82.1%	5	17.9%

By collecting statistics from the evaluation tool, as shown in Table 9.9, one see the following:

1. The MARES method designers have covered all the issues mentioned in the 'criteria and specification', 'Theoretical tools' and 'quantitative data' of the evaluation tool.

2. The MARES method covered about 77% of the issues mentioned in the ‘Fundamental design principles’ class.
3. The MARES method covered about 67% of the issues mentioned in the ‘practical considerations’ class and ‘instrumentalities’ class.

Clearly, the ‘fundamental design principles’ as well as ‘theoretical tools’ need more attention and more improvement.

9.4 Comparisons among the three SPA methods

This section presents several comparisons between the SPA methods discussed in the previous three case studies. Such comparisons provide a view of the strengths and weaknesses of each SPA methods compared to the other methods.

9.4.1 Comparison based on the Bottom-up evaluation method

The comparisons presented in this section are based on the results of the Bottom-up evaluation method. For instance, Figure 9.2 shows the strengths points for each SPA method showing clearly that:

- The $S^{3mAssess}$ method has covered more success evidences related to the assessment method class and supportive tool class than the other two methods;
- While the MARES method has covered more success evidences in the procedure, documentation and users classes than the other two methods.

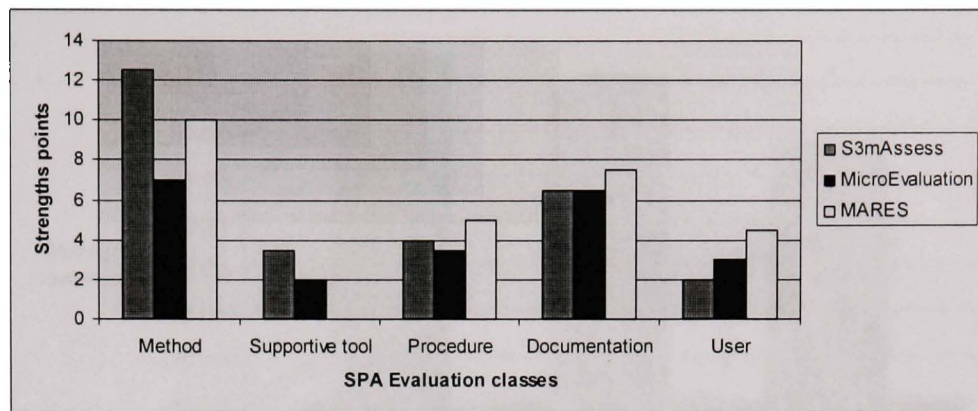


Figure 9.2 Strengths comparisons between the three SPA methods.

Figure 9.3 shows the weaknesses comparisons among the three SPA methods. One observes that the Micro-evaluation method has covered the least number of evidences related to the ‘method’ class and in the ‘documentation’ class. MARES did not cover any of the evidences related to the ‘supportive tool’ class, and finally, the $S^{3mAssess}$ method covered the least number of evidences related to the ‘users’ and ‘documentation’ classes.

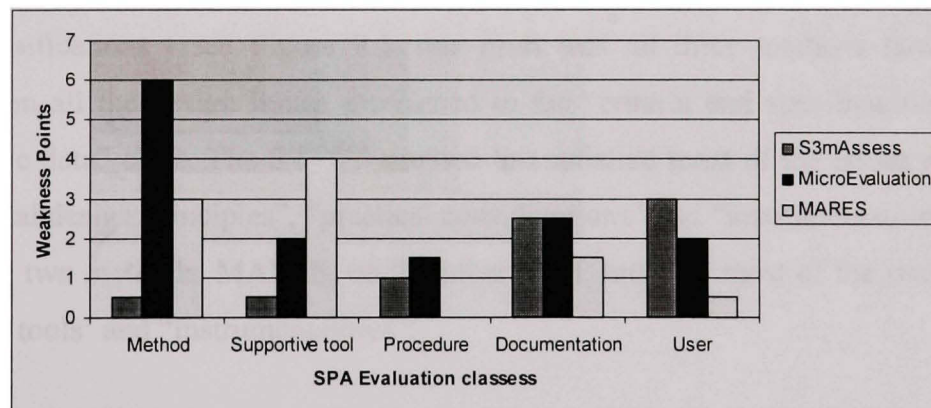


Figure 9.3 Weaknesses comparisons among the three SPA methods.

Overall, the $S^{3mAssess}$ method has more contribution in satisfying the set of success evidences addressed in this thesis, followed by the MARES method and the Micro-evaluation method - see Figure 9.4.

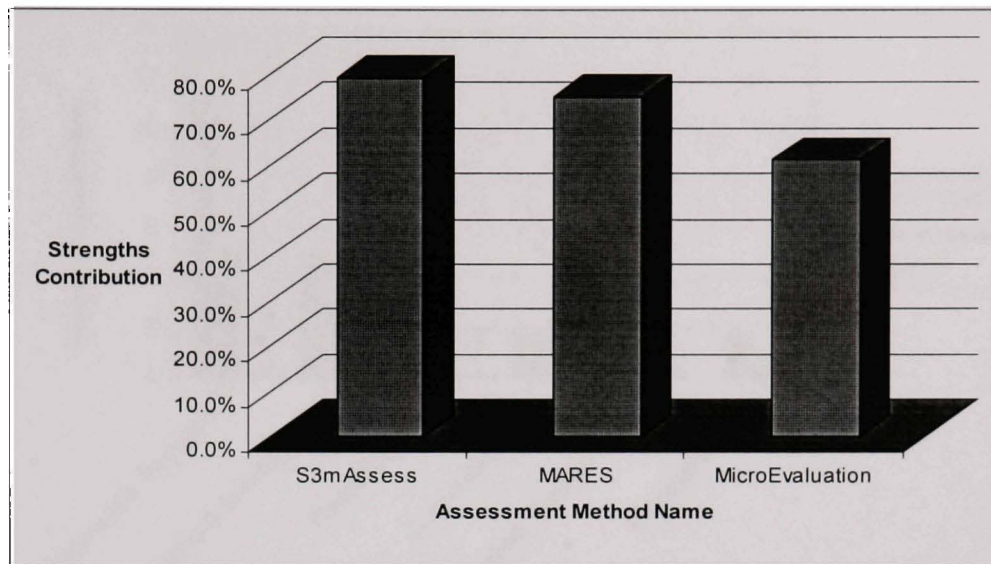


Figure 9.4 Comparison of strength contribution among the three SPA methods.

9.4.2 Comparison based on the Top-up evaluation method

Through a comparison among the three methods based on their alignment with engineering design classifications - see Figure 9.5, one finds that all three methods have taken into consideration all the design issues mentioned in the “criteria and specifications” class and “quantitative data” class. The $S^{3mAssess}$ method has satisfied most of the issues related to the “fundamental design principles”, “practical considerations” and “instrumentalities” compared to the other two methods. MARES, on the other hand, satisfied most of the issues related to ‘theoretical tools’ and “instrumentalities”.

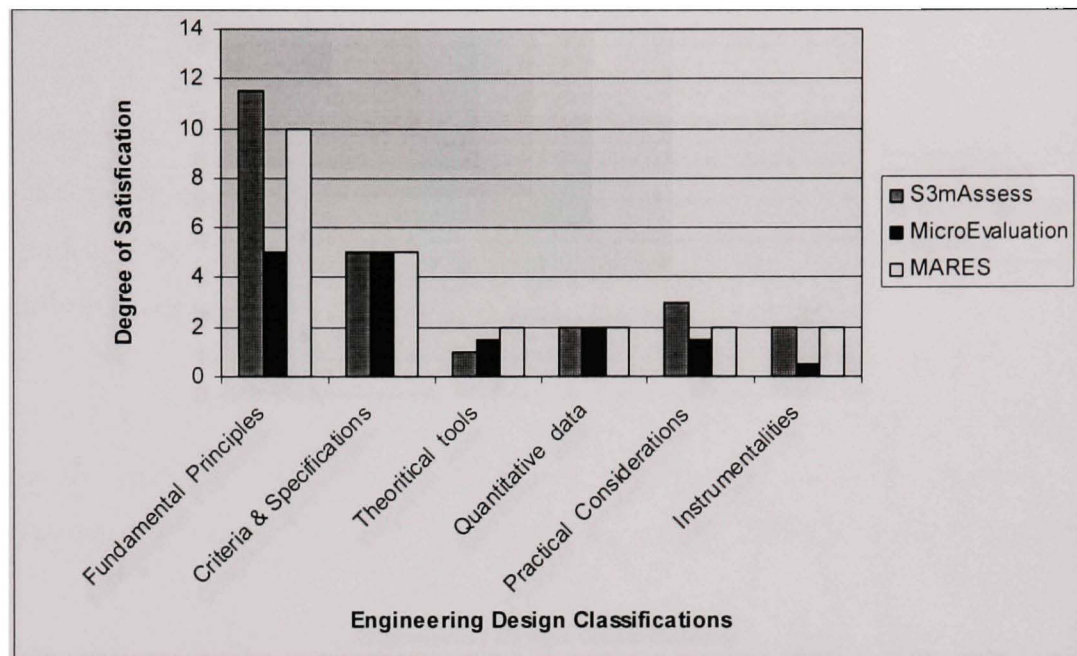


Figure 9.5 Comparison among satisfied design issues of the three SPA methods based on the top-down approach.

Regarding the unsatisfied design issues - see Figure 9.6, the Micro-evaluation method satisfied the least number of issues related tot the “fundamental design principles”, “practical considerations” and “instrumentalities”. While the $S^{3mAssess}$ method satisfied the least number of issues related to the “theoretical tools”.

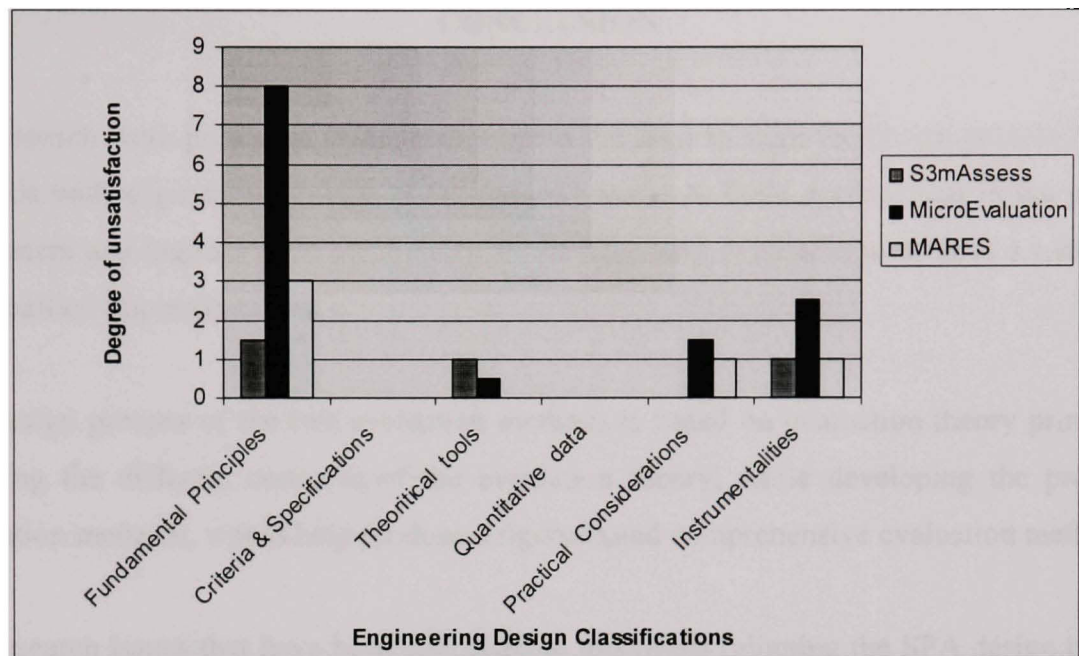


Figure 9.6 Comparison among unsatisfied design issues of the three SPA methods based on the top-down approach.

CONCLUSION

The research work presented in this thesis opens the door to align the design process of SPA methods with engineering design principles and works to build a consensus in the process assessment and improvement community on the necessary evidences to achieve a successful SPA method implementation.

The design process of the two evaluation methods is based on evaluation theory principles. Defining the different concepts of the evaluation theory, while developing the proposed evaluation methods, would help produce a rigorous and comprehensive evaluation method.

The research issues that have been addressed in this thesis (aligning the SPA design process with the engineering design principles and building a consensus on the evidences necessary for successful implementation of SPA methods) have improved the understanding and experience in the design and implementation of lightweight SPA methods. As a result of this research, several contributions have been made, as discussed in the section ‘Contributions of the research’.

The main limitation of the proposed evaluation methods is the level of subjectivity in the rating process. The answers to the questions, mainly the Top-down evaluation methods, are open-ended. The problem of subjectivity can be reduced by adding more specific evidences as discussed in the ‘Future work’ section.

Contributions of the research

1. Identification of comparison criteria that provide useful and informative data suitable for comparison purposes, rather than evaluation between different lightweight SPA methods. The comparison criteria are based on what is available in the literature and few other proposed criteria found suitable and informative to compare between lightweight SPA methods.

2. Identification of evidences found necessary to conduct a successful assessment method. The success evidences are published as success factors, requirements, observations and lessons learned.
3. Application of Vincenti's classifications for engineering design to the design process of the SPA methods. Implementing Vincenti's classifications, generated as a result of the analysis of more than 50 years of experience in the aeronautical engineering design, could enhance the alignment of the SPA methods design with engineering design principles.
4. Evaluation methods of the success of lightweight SPA methods as well as of the design of the lightweight SPA methods.

A number of outcomes of this thesis have been published / submitted in the following journals or conferences.

- Published:
 1. Zarour, M., Desharnais J.-M., and A. Abran, A Framework to Compare Software Process Assessment Methods Dedicated to Small and Very Small Organizations, in International Conference on Software Quality - ICSQ'07. 2007: Denver, CO, USA.
 2. Desharnais, J.M., M. Zarour, and A. April. Very Small Enterprises (VSE) Quality Process Assessment. in 3rd International Workshop on Quality of Information and Communication Technologies. 2007. Havana - Cuba.
 3. Desharnais J. M., C. Y. Laporte, A. Stambollian, M. Zarour, N. Habra, and A. Renault, "Initiating Software Process Improvement with a light model for Small Enterprise: Our Experience," in 3rd International Workshop on Quality of Information and Communication Technologies, Havana - Cuba, 2007.

Future work

As mentioned previously, the research work presented in this thesis opens the door to further work to improve understanding and experience in designing and implementing SPA methods. Light has also been shed on different approaches to design evaluation methods, based on

evaluation theory, systematic literature review and engineering viewpoint, which can be used in other contexts.

Accordingly, the future work that can be pursued based on the results and methodologies used in this thesis can be summarized as follows:

1. Improving the evaluation methods proposed in this thesis: Further improvements to the proposed evaluation methods are possible as experience grows with use. The possible work to be done in the future includes:
 - a. The bottom-up evaluation method can be improved by adding a new level of specific evidences, as shown in Figure 7.5. The evidences in the added level will make the evaluation method more precise and would reduce the subjectivity in the evaluation.
 - b. The proposed evaluation methods have tested through three case studies. The collaboration with more researchers and practitioners to evaluate other lightweight SPA methods and get their feedbacks would help improve the proposed evaluation method. More work is needed to evaluate other lightweight SPA methods.
2. The ISO group WG24 is currently working to produce software lifecycle profiles for VSE organizations. The third part of their five parts standard draft, ISO 29110, is concerned with presenting an assessment guide. This guide addresses the requirements for performing assessment in VSE organizations. Accordingly, the findings of this research work such as evidences and guidelines for designing assessment methods, which have been used to build the proposed evaluation methods, would provide useful input for ISO 29110-3 and provide the designers as well as assessors with guidelines in designing and conducting assessments.
3. ISO 15504 standard (2003) addresses the process assessment and the application of process assessment for improvement and capability determination. “It defines the minimum set of requirements for performing an assessment that will ensure assessment results are objective, impartial, consistent, repeatable and representative of the assessed processes” (ISO 15504-2 2003).

As stated by ISO 15504-2:2003 concerning part 2 of the standard: “It is primarily addressed to the competent assessor and other stakeholders, such as the sponsor of the assessment, who need to be assured that the requirements of this international standard have been met. It will also be of value to developers of assessment methods and of tools to support an assessment”. The standard is written to benefit the assessors and stakeholders primarily rather than the SPA methods designers. However, as can be seen from the literature review, the SPA methods design based on ISO 15504 requirements, used the ISO documents as their primary references. Hence, the needs of these designers as users of this standard should be stated explicitly in the standard.

Accordingly, the SPA methods design issues including guidelines and evidences for successful implementation of assessment methods should also be included in the ISO 15504 standards mainly in parts 2 and 3, requiring that these two parts of the ISO 15504 be rewritten and reorganized in a way that focuses both in designing and conducting assessments. The different guidelines and success evidences for designing and implementing SPA methods presented in this thesis can be generalized and used in the new modified versions of ISO 15504-2 and 15504-3.

4. The design approach implemented in this thesis to develop the proposed evaluation methods can be used to design other models or processes in other contexts than SPA. Moreover, the approach can be used to ‘evaluate’ other types of Assessment methods, including:
 - a. In process assessment test, design, measurement.
 - b. In software product assessment quality, security.

ANNEX I

THE EVALUATION TOOL DEVELOPED FOR THE EVALUATION METHOD BASED ON ENGINEERING DESIGN PRINCIPLES

SPA Process Evaluation Criteria based on Fundamental Design Principles		
No.	Criteria	Answer
SPA-FQ1	What process reference model is the SPA method based on?	
SPA-FQ2	What process assessment model is the SPA method based on?	
SPA-FQ3	Does the SPA method define the business need before the assessment?	
SPA-FQ4	Does the SPA method make use of previous assessment reports?	
SPA-FQ5	Does the SPA method refer to the organizational documents and reports while preparing for the assessment?	
SPA-FQ6	Does the assessment method make use of assessment tools through different phases of the assessment?	
SPA-FQ7	Does the assessment process produce a pre-assessment questionnaire to collect information that helps structure the on-site interview?	
SPA-FQ8	Does the assessment method produce assessment initiation file?	
SPA-FQ9	Does the assessment method produce assessment plan?	
SPA-FQ10	Does the assessment method track evidences of process performance and capability?	
SPA-FQ11	Does the assessment process produce an assessment report?	
SPA-FQ12	Does the assessment process produce assessment record?	
SPA-FQ13	Does the assessment process produce assessor record?	

SPA Process Evaluation Criteria based on Criteria and Specifications		
No.	Criteria	Answer
SPA-CQ1	What is the number of processes to be assessed?	
SPA-CQ2	What are the processes to be assessed?	
SPA-CQ3	What are the criteria for assessing the process (how the processes are rated, please provide details?)	
SPA-CQ4	What is the scale and its limits used to assess the process?	
SPA-CQ5	What is the scale and its limits used to assess the organization?	
SPA Process Evaluation Criteria Based on Theoretical Tools		
No.	Criteria	Answer
SPA-TQ1	What theoretical tools are used to select the processes to be assessed?	
SPA-TQ2	What theoretical tools are used to define the rating process?	
SPA Process Evaluation Criteria based on Quantitative Data		
No.	Criteria	Answer
SPA-QQ1	What data you use to determine the scale for each process and obtain the results?	
SPA-QQ2	What data you use to determine the scale for the organization?	
SPA Process Evaluation Criteria Based on Practical Considerations		
No.	Criteria	Answer
SPA-PQ1	How are selected the processes to be assessed?	
SPA-PQ2	What is the target scaling level for the organization?	
SPA-PQ3	Does the assessment method build an action plan at the end of the assessment?	

SPA Process Evaluation Criteria Based on Instrumentalities		
No.	Criteria	Answer
SPA-IQ1	What are the steps and sub-steps of the assessment method that should be defined during the assessment design process?	
SPA-IQ2	What judgments do the designer takes related to which processes to be assessed?	
SPA-IQ3	What procedure is used to improve the design of the SPA method?	

ANNEX II

SUMMARY OF THE PUBLICATIONS REVIEWED IN THIS THESIS AND FOUND TO DISCUSS SPECIFIC SUCCESS EVIDENCES RELATED TO THE SPA METHODS

References	Description
(Loon, Cass et al. 2004)	This book looks at the practical aspects of process assessment and their use. It is a comparison book to process assessment and ISO 15504. The book discusses the major business reasons for process assessment, provides useful figures on the costs of poor quality, and the benefits of the improvements for organizations. It also provides comprehensive programmes and agile methodologies for process improvement.
(CMMI-SUT 2006)	The Appraisal Requirements for CMMI (ARC) V1.2 defines the requirements considered essential to appraisal methods intended for use with Capability Maturity Model Integration-CMMI models. In addition, a set of appraisal classes is defined, based on typical applications of appraisal methods. These classes are intended primarily for developers of appraisal methods to use with CMMI capability models in the context of the CMMI Product Suite. This document defines the requirements for such methods, but not necessarily the conditions or constraints under which they might be applied.
(Humphrey and Kitson 1987)	This report describes the methodology developed by the Software Engineering Institute (SEI) for conducting SEI-assisted assessments of software engineering capability. The assessment methodology has five phases: 1) selecting the candidate organization, 2) preparing for the assessment, 3) conducting the assessment, 4) communicating final assessment findings and action recommendations, and 5) post-assessment follow-up activities.
(Wangenheim, Varkoi et al. 2005)	This paper presents a flexible approach to efficient process assessments in small organizations. Flexibility requires a continuous assessment model, so that the scope of process improvement and assessment can be defined on the basis of the prioritized needs of an organization. Based on their experiences, the authors showed that ISO 15504 can also be applied with success in small software organizations. This paper presents how the assessments were run and lessons learned on applying 15504 in small organization.
(Wangenheim, Varkoi et al. 2006)	This paper presents a flexible approach to efficient process assessments in small organizations. Flexibility requires a continuous assessment model, so that the scope of process improvement and assessment can be defined on the basis of the prioritized needs of an organization. Based on their experiences, the authors showed that ISO 15504 can also be applied with success in small software organizations. This paper presents how the assessments were run and lessons learned on applying 15504 in small organization.

References	Description
(Anacleto, Wangenheim et al. 2004)	This paper describes some experience gained from applying ISO 15504 for software process assessment focusing on process improvement in 4 small Brazilian companies. Cost and benefit related to the assessment are also presented.
(Anacleto, Wangenheim et al. 2004)	This article discusses the assessment method developed by the authors as part of the research project 15504MPE with the objective of process improvement adapted on the small Brazilian software organizations. Various assessments have been performed in small organizations using the developed method.
(Cater-Steel, Toleman et al. 2006)	This paper presents an evaluation of a program in which low-rigour, one-day SPI assessments were offered at no cost to 22 small Australian software development firms. The assessment model was based on ISO/IEC 15504(SPICE). About 12 months after the assessment, the firms were contacted to arrange a follow-up meeting to determine the extent to which they had implemented the recommendations from the assessment. Comparison of the process capability levels at the time of assessment and the follow-up meetings revealed that the process improvement program was effective in improving the process capability of 15 of these small software development firms. The article also talked about possible improvements to the assessment method used.
(Cater-Steel 2002)	This research investigates the adoption of SPI initiatives by four small software development firms. These four firms participated in a process improvement program which was sponsored by Software Engineering Australia (SEA) (Queensland). The assessment method was based on SPICE (ISO/IEC 15504) and included an initial assessment, recommendations, and a follow-up meeting for each firm. The article summarizes the improvements realized and considers the critical success factors relating to SPI adoption for small firms.
(Fabbrini, Fantini et al. 2003)	This article discuss the requirements that a SPICE assessment must satisfy to fulfill the needs of different stakeholders, then the requirements an automatic tools has to satisfy in order to be able to provide support in the different phases of the assessment.
(Kautz 1998)	This paper discusses how and under which circumstances software process improvement can be rewarding for small software companies. It describes the course of a process improvement project which was jointly carried out by three small companies and discusses the critical success factors for SPI in small companies.
(McCaffery, Taylor et al. 2007)	The authors in this article have created an assessment method called Adept. Adept method combines two process assessment methods, one plan-based and one agile-based, without dictating which one to use. This makes the lightweight approach attractive for small software companies.

References	Description
(Komiya, Sunazuka et al. 2001)	This paper proposes a framework for Software Process Assessment and Improvement (SPAIM) and its performance measurement. The main purpose of this framework is to make SPAIM-related technologies adaptable to the features of an assessed organization such as organizational goals, future products, etc. The key to enact this framework is construction of SPAIM libraries containing various technologies for assessing, improving, and measuring software processes. Then, we can compose a specific SPAIM method adapted to the assessed organization by selecting and customizing technologies included in the libraries. This concept has been developed through more than 10 years of software process improvement experience.
(Grünbacher 1997)	This paper reports on the experiences of an assessment and improvement process in a small software company. The process is based on an initial self-evaluation and following structured interviews in the company. A first application of the approach revealed very promising results.
(Stambollian, Habra et al. 2006)	This paper presents the results obtained by applying the Micro-Evaluation assessment method in several small settings. It summarizes the experience gained and lessons learned by the application of this method.
(Laporte, Desharnais et al. 2005)	In this article the authors present the results of their research project on software process improvement in small organizations in Wallonia-Belgium and Quebec-Canada.
(Habra, Eustache Niyitugabira et al. 1999)	This paper presents a technology transfer experience which aims at supporting the introduction of software process improvement in small businesses, small organizations and/or small projects. The experience is an European interregional collaboration between two university research teams (France and Belgium) and a public technology center (Luxembourg). One of the contributions of this experience is the design of a Software Process Improvement approach particularly adapted to small units on the one hand, and to regional context, on the other hand. The proposed approach is gradual. It is based on three nested evaluation models ranging from an extremely simplified model (the micro-evaluation model) to a complete standard model which is a version of SPICE. The intermediate model, called the mini-evaluation model, can be viewed as a tailoring of SPICE and can be used by itself as a definitive model by small businesses and small organizations.

References	Description
(Dyba and Moe 1999)	This paper presents a critique of the global "best practice" approach to software process assessment and improvement, focusing on the necessity to explore the contingencies of individual software organizations. Furthermore, the paper presents some of the author experiences in using tailor made assessments based on a participative approach to focus software process improvement activities in Norwegian software companies.
(Ekdahl and Larsson 2006)	Critical to any successful performance improvement initiative is to achieve a state of continuous or institutionalized improvement. Some improvement can happen quickly, but long-term improvement is typically a matter of sustaining focus. In this paper the use of internal appraisals as a means for sustaining improvement focus is discussed. Experiences and lessons learnt, as well as some of the specifics to the internal appraisals of an organization called ABB's is presented.
(Wiegers and Sturzenberger 2000)	This paper develops a CMM-based modular mini-assessment method called MMA to meet the diverse needs of software projects undertaking process improvement efforts at Kodak. The authors said that their method's components are well defined yet flexible, offering project leaders several options. The results described in this paper illustrate the method's application, benefits, and limitations.
(ISO/IEC 2003)	ISO/IEC 15504 part 3 provides guidance on meeting the minimum set of requirements for performing an assessment contained in ISO/IEC 15504-2. It provides an overview of process assessment and interprets the requirements.
(Cater-Steel 2001)	This paper presents findings from the RAPID process improvement program as experienced by four small software companies in Australia. For each company, the initial assessment findings are presented in this paper followed by the outcome of the recommendations. The discussion which follows this presentation compares the capability of the four companies and links the findings from this case study to existing literature related to software process improvement success and failure.
(Choi, Lee et al. 2005)	In this paper, the authors suggest process a management system which is needed to assist SPI activities. To enhance the quality of software product, the proposed system supports internal assessors to conduct the gap analysis compliant with ISO IS 15504. Later, it has evolved as a process mentoring tool for developers, as well as SPI tool for software development organizations.

References	Description
(Chen, Sorenson et al. 2007)	This paper introduces a web-based prototype system, Continuous SPA, on continuous assessing and monitoring of software process, and performs a practical study in one process area: project management. The study results are presented and show that features such as global management, well-defined responsibility and visualization can be integrated in process assessment to help improve the software process management.
(Salvaneschi, Grasso et al. 2006)	This article presents the experience gained and the lessons learned during the process assessment and improvement of the IT department of a large retail company. The business motivations and the main approaches and results of the assessment process are presented and discussed. The assessment led to an improvement in the project that is currently running. The article discussed some aspects that may be useful for assessing and improving similar types of organizations. Among them is the role of the global service provider in IT departments, the relevance of communication and knowledge management processes and the need for assessment tailoring and people involvement.
(Lok and Walker 1997)	This paper describes the development of an automated tool to support the conduct of assessments in Phase 2 of the trialling period using the embedded model described in Part 5 of ISO 15504. The development process and product quality objectives and their validation are described. The impact of user feedback on the evolution of the tool is reviewed. Experience in the application of the tool for conducting assessments is also presented.
(Pettersson, Ivarsson et al. 2008)	This paper presents a step-by-step guide to process assessment and improvement planning using improvement framework utilizing light weight assessment and improvement planning iFLAP, aimed at practitioners undertaking SPI initiatives. In addition to the guide itself the industrial application of iFLAP is shown through two industrial cases. iFLAP is a packaged improvement framework, containing both assessment and improvement planning capabilities, explicitly developed to be light weight in nature. Assessment is performed by eliciting improvements issues based on the organization's experience and knowledge.
(Santos, Montoni et al. 2007; Santos, Montoni et al. 2007)	In order to support the SPI activities, a SPI deployment strategy named SPI-KM that is supported by Knowledge Management has been developed in this paper based on international and national reference models and standards. This paper presents the SPI-KM strategy and the results of an empirical study executed aiming to characterize the SPI initiatives that employed it. The study findings are presented as lessons learned and their applications are discussed in different organizations.

References	Description
(Cater-Steel 2004)	This is a research study that aims at understanding the extent of software development practices currently in use and to evaluate the effectiveness of assessment-based software process improvement initiatives for small firms. This study presents the results of assessment process for 22 organizations and the associated improvement programs carried out by these organizations.

ANNEX III

REQUIREMENTS FOR CLASS-C CMMI APPRAISAL METHODS

Requirements	Description	Class C
Responsibilities		
Appraisal Sponsor	<p>The method shall define the responsibilities of the sponsor, which includes at minimum:</p> <ul style="list-style-type: none"> a. Verify that the appraisal team leader has the appropriate experience, knowledge, and skills to take responsibility for and lead the appraisal. b. Ensure that the appropriate organizational units or subunits (e.g., projects, functional units) participate in the appraisal. c. Support appraisal method provisions for ensuring non-attribution to appraisal participants. d. Ensure that resources are made available to conduct the appraisal. e. Review and approve the appraisal input prior to the beginning of data collection by the appraisal team. 	yes
Appraisal Team Leader	<p>The method shall define the responsibilities of the team leader, which shall includes at minimum:</p> <ul style="list-style-type: none"> a. Ensure that the appraisal is conducted in accordance with the method's documented process. b. Confirm the sponsor's commitment to proceed with the appraisal. c. Ensure that appraisal participants are briefed on the purpose, scope, and approach of the appraisal. d. Ensure that all appraisal team members have the appropriate experience, knowledge, and skills in the appraisal reference model and appraisal method; the necessary competence to use instruments or tools chosen to support the appraisal; and access to documented guidance on how to perform the defined appraisal activities. e. Verify and document that the appraisal method requirements have been met. f. Confirm delivery of appraisal results to the sponsor. 	yes

Requirements	Description	Class C
Appraisal Method Documentation		
Documentation of method	<p>The method shall be documented and, at a minimum, include:</p> <ul style="list-style-type: none"> a. Identification of the CMMI models (version, discipline, and representation [staged or continuous]) with which the method can be used. b. Identification of the ARC version upon which the appraisal method is based. c. Identification of which CMMI appraisal requirements are satisfied by the method, along with the CMMI appraisal method class membership (if applicable). d. Activity descriptions, artifacts, and guidance that implement each of the appraisal requirements. 	partial (a-d only)
Guidance for identifying appraisal purpose and objectives	<p>The method documentation shall provide guidance for</p> <ul style="list-style-type: none"> a. Identifying an appraisal's purpose, objectives, and constraints. b. Determining the suitability of the appraisal method relative to the appraisal's purpose, objectives, and constraints. 	yes
Guidance for CMMI model scope	<p>The method documentation shall provide guidance for identifying the scope of the CMMI model(s) to be used for the appraisal:</p> <ul style="list-style-type: none"> a. Process areas to be investigated (continuous and staged representations). b. Capability levels to be investigated for each process area (continuous representation). 	yes
Guidance for identifying organizational unit	<p>The method documentation shall provide guidance for identifying the organizational unit to be appraised:</p> <ul style="list-style-type: none"> a. The sponsor of the appraisal and the sponsor's relationship to the organizational unit being appraised. b. Projects within the organizational unit that will participate. c. Functional elements of the organizational unit that will participate. d. Names and affiliations (organizational units) of participants in the appraisal activities. 	yes

Requirements	Description	Class C
Appraisal Method Documentation		
Guidance for team member selection	<p>The method documentation shall provide guidance for selecting appraisal team members and criteria for qualification, including:</p> <ul style="list-style-type: none"> a. Technical experience (discipline-specific). b. Management experience. c. Experience, knowledge, and skills in the appraisal reference model and appraisal method. 	yes
Guidance for team leader qualification criteria	<p>The method documentation shall provide guidance for an appraisal team leader's qualification criteria, including</p> <ul style="list-style-type: none"> a. Training and experience using the appraisal reference model. b. Training and experience using the appraisal method. c. Experience in delivering training, managing teams, facilitating group discussions, and making presentations. 	yes
Guidance for size of team	The method documentation shall provide guidance for determining the appropriate size of the appraisal team. For Class C appraisals, one member.	yes
Guidance for team member roles and responsibilities	The method documentation shall provide guidance on the roles and responsibilities of appraisal team members.	yes
Guidance for appraisal sponsor responsibilities	The method documentation shall provide guidance addressing the responsibilities of the appraisal sponsor.	yes
Guidance for team leader responsibilities	The method documentation shall provide guidance addressing the responsibilities of the appraisal team leader.	yes
Guidance for estimating appraisal resources	The method documentation shall provide guidance for estimating the resources required to conduct the appraisal (including the amount of time required to conduct an appraisal).	yes
Guidance for logistics	The method documentation shall provide guidance for appraisal logistics.	yes
Guidance for collecting and mapping data to appraisal reference model	The method documentation shall provide guidance for collecting relevant data on the organizational unit and associating the data to the specific and generic practices of the appraisal reference model.	yes

Requirements	Description	Class C
Guidance for creation of findings	The method documentation shall provide guidance for creating findings, including both strengths and weaknesses relative to the appraisal reference model.	yes
Guidance for assuring confidentiality and non-attribution	The method documentation shall provide guidance for protecting the confidentiality of appraisal data and ensuring non-attribution of data contributed by appraisal participants.	yes
Guidance for appraisal record	<p>The method documentation shall provide guidance for (1) recording traceability between the data collected during the appraisal and the findings and/or ratings, (2) the retention and safekeeping of appraisal records, and (3) compiling and maintaining an appraisal record that supports the appraisal team's findings and/or ratings and that contains the following minimum content:</p> <ul style="list-style-type: none"> a. Dates of appraisal. b. Appraisal input. c. Objective evidence or identification thereof, sufficient to substantiate goal rating judgments. d. Identification of appraisal method (and version) used, along with any tailoring options. e. Findings. 	partial (a,b,d,e only)
Planning and Preparing for the Appraisal		
Preparation of participants	<p>The method shall provide for the preparation of appraisal participants by addressing, at a minimum:</p> <ul style="list-style-type: none"> a. The purpose of the appraisal. b. The scope of the appraisal. c. The appraisal approach. d. The roles and responsibilities of participants in the appraisal. e. The schedule of appraisal activities. 	yes
Development of appraisal input	The method shall provide for the development of the appraisal input prior to the beginning of data collection by the appraisal team.	yes

Requirements	Description	Class C
Planning and Preparing for the Appraisal		
Content of appraisal input	<p>At a minimum, the appraisal input shall specify</p> <ol style="list-style-type: none"> a. The identity of the sponsor of the appraisal, and the sponsor's relationship to the organizational unit being appraised. b. The appraisal purpose, including alignment with business objectives. c. The appraisal reference model scope, including: <ol style="list-style-type: none"> 1. The process areas to be investigated within the organizational unit. 2. The highest maturity level and/or capability level to be investigated for each process area within the appraisal scope. d. The organizational unit that is the subject of the appraisal. e. The process context, which, at a minimum, shall include: <ol style="list-style-type: none"> 1. The size of the organizational unit. 2. The demographics of the organizational unit. 3. The application domain of the products or services of the organizational unit. 4. The size, criticality, and complexity of the products or services. f. The appraisal constraints, which, at a minimum, shall include: <ol style="list-style-type: none"> 1. Availability of key resources (e.g., staffing, funding, tools, facilities). 2. Schedule constraints. 3. The maximum amount of time to be used for the appraisal. 4. Specific process areas or organizational entities to be excluded from the appraisal. 5. The minimum, maximum, or specific sample size or coverage that is desired for the appraisal. 6. The ownership of the appraisal outputs and any restrictions on their use. 7. Controls on information resulting from a confidentiality agreement. 8. Non-attribution of appraisal data to associated sources. g. The identity of the CMMI models used, including the version, discipline, and representation (staged or continuous). h. The criteria for experience, knowledge, and skills of the appraisal team leader who is responsible for the appraisal. i. The identity and affiliation of the appraisal team members, including the appraisal team leader, with their specific appraisal responsibilities. 	partial (all except e.5)

Requirements	Description	Class C
Planning and Preparing for the Appraisal		
	<ul style="list-style-type: none">j. The identity (name and organizational affiliation) of appraisal participants and support staff, with specific responsibilities for the appraisal.k. Any additional information to be collected during the appraisal to support achievement of the appraisal objectives.l. A description of the planned appraisal outputs, including ratings to be generated (process areas, maturity level).m. Anticipated follow-on activities (e.g., reports, appraisal action plans, re-appraisal).n. Planned tailoring of the appraisal method and associated tradeoffs, including the sample size or coverage of the organizational unit.	
Sponsor approval of appraisal input	The method shall require that the appraisal input, and any changes to the appraisal input, shall be agreed to by the sponsor (or the delegated authority) and documented in the appraisal record.	yes
Development of appraisal plan	<p>The method shall require the development of an appraisal plan that, at a minimum, specifies:</p> <ul style="list-style-type: none">a. The appraisal input.b. The activities to be performed in conducting the appraisal.c. Resources and schedule assigned to appraisal activities.d. Appraisal logistics.e. Mitigation steps to address risks associated with appraisal execution.	partial (a-e only)
Appraisal Data Collection		
Data from interviews	The method shall collect data by conducting interviews (e.g., with project leaders, managers, practitioners).	At least one type of objective evidence
Data from documents	The method shall collect data by reviewing documentation (e.g., organizational policies, instruments, project procedures, and implementation-level work products).	
Data Consolidation and Validation		
Consensus of team members	The method shall require appraisal team consensus when teams are involved in Decision related to determining the validity of findings and establishing ratings.	Yes

Requirements	Description	Class C
Data Consolidation and Validation		
Accuracy of findings	<p>The method shall require a mechanism for consolidating the data collected during an appraisal into accurate findings according to the following criteria:</p> <ol style="list-style-type: none"> The finding was derived from objective evidence seen or heard during data collection sessions. The finding is clearly worded, phrased without attribution, and expressed in terminology used at the organizational unit. Objective evidence supporting the finding is traceable to the project or organizational unit. The finding is relevant to the appraisal reference model and can be associated with a specific model component. 	Yes
Validation of findings	<p>The method shall require a mechanism for verifying findings according to the following criteria:</p> <ol style="list-style-type: none"> The finding is based on corroborated objective evidence. The finding is consistent with other verified findings. (Verified findings cannot be both true and mutually inconsistent; in aggregate, they constitute a set of truths about the organizational unit that must be consistent). 	optional
Corroboration of objective evidence	<p>The method shall require the following minimum set of criteria to be satisfied in order for objective evidence to be considered “corroborated”:</p> <ol style="list-style-type: none"> The objective evidence is obtained from at least two different sources. At least one of the two sources must reflect work actually being done (e.g., process area implementation). 	optional
Sufficiency of data	The method shall require a mechanism for determining that sufficient data has been collected to cover the scope of the appraisal.	Optional
Preliminary findings preparation	The method shall require a mechanism for consolidating objective evidence into preliminary findings of strengths and weaknesses relative to the appraisal reference model.	Optional
Preliminary findings presentations	The method shall require that appraisal participants be presented with the preliminary findings in order to solicit their responses for validation of the findings’ accuracy and clarity.	Optional

Requirements	Description	Class C
Reporting Results		
Report results to sponsor and appraised organization	The method shall require documenting and reporting the appraisal findings and/or ratings to the appraisal sponsor and to the appraised organization.	Yes
Appraisal results to CMMI Steward	The method shall require the submission of appraisal data required by the CMMI Steward for the purpose of reporting aggregated appraisal information to the constituent community.	Optional
Retention of appraisal record	The method shall require that the appraisal record be provided to the appraisal sponsor for retention.	Yes

ANNEX IV

SUMMARY OF THE REQUIREMENTS FOR PERFORMING ASSESSMENTS AS STATED IN ISO 15504-3

Activity	Sub-activities	Requirements
Assessment Activities	Planning: Develop and document assessment plan	<ol style="list-style-type: none">1. The required inputs defined in this part of ISO/IEC 15504.2. The activities to be performed in conducting the assessment.3. The resources and schedule assigned to these activities.4. The identity and defined responsibilities of the participants in the assessment.5. The criteria to verify that the requirements of this International Standard have been met.6. A description of the planned assessment outputs.
	Data Collection	<ol style="list-style-type: none">1. The strategy and techniques for the selection, collection, analysis of data and justification of the ratings shall be explicitly identified and shall be demonstrable.2. Correspondence shall be established between the organizational unit's processes, specified in the assessment scope, and the elements in the Process Assessment Model.3. Each process identified in the assessment scope shall be assessed on the basis of objective evidence.4. The objective evidence gathered for each attribute for each process assessed shall be sufficient to meet the assessment purpose and scope.5. The identification of the objective evidence gathered shall be recorded and maintained to provide the basis for verification of the ratings.

Activity	Sub-activities	Requirements
Assessment Activities (Continued)	Data Validation	<ol style="list-style-type: none"> 1. Confirm that the evidence collected is objective. 2. Ensure that the objective evidence is sufficient and representative to cover the scope and purpose of the assessment. 3. Ensure that the data as a whole is consistent.
	Process attribute rating	<ol style="list-style-type: none"> 1. The set of process attribute ratings shall be recorded as the process profile for the defined organizational unit. 2. During the assessment, the defined set of assessment indicators in the Process Assessment Model shall be used to support the assessor's judgement in rating process attributes in order to provide the basis for repeatability across assessments. 3. The decision-making process that is used to derive rating judgements shall be recorded. 4. Traceability shall be maintained between an attribute rating and the objective evidence used in determining that rating. 5. For each process attribute rated, the relationship between the indicators and the objective evidence shall be recorded.
	Reporting	A final report should be prepared and delivered to the organizations participants.
Roles and responsibilities	Sponsor Responsibilities	<ol style="list-style-type: none"> 1. Verify that the individual who is to take responsibility for conformity of the assessment is a competent assessor. 2. Ensure that resources are made available to conduct the assessment. 3. Ensure that the assessment team has access to the relevant resources.

Activity	Sub-activities	Requirements
	Competent assessor responsibilities	<ol style="list-style-type: none"> 1. Confirm the sponsor's commitment to proceed with the assessment. 2. Ensure that the assessment is conducted in accordance with the requirements of this part of ISO/IEC 15504. 3. Ensure that participants in the assessment are briefed on the purpose, scope and approach of the assessment. 4. Ensure that all members of the assessment team have knowledge and skills appropriate to their roles. 5. Ensure that all members of the assessment team have access to appropriate documented guidance on how to perform the defined assessment activities. 6. Ensure that the assessment team has the competencies to use the tools chosen to support the assessment. 7. Confirm receipt of the assessment result deliverables by the sponsor. 8. On completion of the assessment, verify and document the extent of conformance of the assessment to ISO/IEC 15504 (see also clause 7.4).
	Assessor responsibilities	<ol style="list-style-type: none"> 1. Carry out assigned activities associated with the assessment, e.g. detailed planning, data collection, data validation and reporting. 2. Rate the process attributes.

Activity	Sub-activities	Requirements
Define assessment input		<ol style="list-style-type: none"> 1. The identity of the sponsor of the assessment and the sponsor's relationship to the organizational unit being assessed. 2. The assessment purpose. 3. The assessment scope including: <ol style="list-style-type: none"> a. The processes to be investigated within the organizational unit. b. The highest capability level to be investigated for each individual process within the assessment scope. c. The organizational unit that deploys the processes. d. The context which includes: <ul style="list-style-type: none"> • The size of the organizational unit; • The application domain of the products or services of the organizational unit; • Key characteristics (e.g. size, criticality, complexity and quality) of the products or services of the organizational unit. 4. The assessment approach. 5. The assessment constraints considering, at minimum: <ol style="list-style-type: none"> a. Availability of key resources. b. The maximum duration of the assessment. c. Specific processes or organisational units to be excluded from the assessment. d. The quantity and type of objective evidence to be examined in the assessment. e. The ownership of the assessment outputs and any restrictions on their use. f. Controls on information resulting from a confidentiality agreement. 6. The identity of the Process Assessment Model (including the identity of the Process Reference Model(s) used) that meets the requirements defined in clause 6.3 of the ISO 15504-3. 7. The identity of the competent assessor. 8. The criteria for competence of the assessor who is responsible for the assessment.

Activity	Sub-activities	Requirements
Define assessment input (continued)		<ol style="list-style-type: none"> 9. The identity and roles of assessees, the assessment team and assessment support staff with specific responsibilities for the assessment. 10. Any additional information to be collected during the assessment to support process improvement o process capability determination, e.g. specific data (or metrics) that are needed to quantify the organization's ability to meet a particular business goal (this may also include information detailed a clause 6.3.5 and associated note of the ISO 15504-3).
Recording assessment output		<ol style="list-style-type: none"> 1. The date of the assessment. 2. The assessment input. 3. The identification of the objective evidence gathered. 4. Identification of the documented assessment process. 5. The set of process profiles resulting from the assessment (i.e. one profile for each process assessed). 6. The identification of any additional information collected during the assessment as specified in clause .4.2(j) of the ISO 15504-3.

ANNEX V

SUMMARY OF THE EVIDENCES FOUND NECESSARY FOR THE SUCCESS OF THE ASSESSMENT METHODS AND THEIR REFERENCES

ECID: Evidence Class ID, a number from 1-5 (1: Documentation, 2: Method, 3: Tools, 4: Procedures, 5: Users).

EN: Evidence number (sequence number for evidences in each evidence class)

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
Book	2004	Loon	SPA/ SPI			Success Factor	5	4	Commitment
						Success Factor	5	5	Motivation of participants
						Success Factor	4	4	Confidentiality so the participants do not feel threats
						Success Factor	5	5	Benefits: the participants should feel the benefits of the assessment
						Success Factor	5	6	Credibility: the sponsor and staff should believe that the assessment will give results
Technical Report	2006	CMMI	SPA			Requirement	1	6	Documentation of method

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	1	1	Guidance for identifying appraisal purpose and objectives
						Requirement	1	2	Guidance for identifying organizational unit
						Requirement	1	3	Guidance for team member selection
						Requirement	1	3	Guidance for team leader qualification criteria
						Requirement	1	3	Guidance for size of team
						Requirement	1	3	Guidance for team member roles and responsibilities
						Requirement	1	3	Guidance for appraisal sponsor responsibilities
						Requirement	1	3	Guidance for team leader responsibilities
						Requirement	1	6	Guidance for estimating appraisal resources
						Requirement	1	1	Guidance for logistics
						Requirement	1	6	Guidance for collecting and mapping data to appraisal reference model

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	1	7	Guidance for creation of findings
						Requirement	1	4	Guidance for assuring confidentiality and non-attribution
						Requirement	1	6	Guidance for appraisal record
						Requirement	2	1	Data from interviews
						Requirement	2	2	Data from documents
						Requirement	2	3	Accuracy of findings
						Requirement	4	3	Report results to sponsor and appraised organization
						Requirement	4	1	Preparation of participants
						Requirement	4	1	Development of appraisal input
						Requirement	4	1	Content of appraisal input
						Requirement	4	1	Sponsor approval of appraisal input

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	4	1	Development of appraisal plan
						Requirement	5	1	Appraisal Sponsor Responsibilities
						Requirement	5	2	Appraisal Team Leader Responsibilities
						Requirement	2	3	Consensus of team members when validating collected data
						Requirement	1	6	Retention of appraisal record
Technical Report	1989	Humphrey	SPA			Requirement	2	7	Action orientation
						Requirement	4	4	Confidentiality
						Requirement	4	5	Non-adversarial attitude
						Requirement	5	3	Senior Management Involvement
EUROSPICE '2005 Conference	2005	Wangenheim	SPA	Small	16	Lessons Learned	1	5	Availability of documents templates
Software Process Improve. and Practice	2006	Wangenheim	SPA	Small	16	Lessons Learned	2	4	Flexible assessment model based on continuous representation

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Lessons Learned	2	4	Focus on principal high-priority processes
						Lessons Learned	2	5	Coverage of the process reference model
						Lessons Learned	2	1	Data collection based on group interviews
						Lessons Learned	2	6	Identification of risks and improvement opportunities
						Lessons Learned	3	1	Availability of tool support
SPICE Conference	2004	A. Anacleto	SPA	Small	1	Requirement	2	10	Low assessment cost
						Requirement	2	12	Reliable assessment results
						Requirement	1	6	Detailed description of the assessment process including explicit guidance for its application in practice
						Requirement	2	4	Flexible assessment method based on continuous representation
						Requirement	1	6	Detailed description of the assessment model including measurement framework and process reference model
						Requirement	2	6	Support for identification of risks and improvement suggestions

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	1	6	Support for the description of the a high-level model of the assessed process
						Requirement	2	9	Conformity with ISO 15504
						Requirement	5	1	Not requiring any specific software engineering knowledge form company representative
						Requirement	3	1	Supported by software tool covering complete assessment process
						Requirement	2	9	Integrated in an assessment methodology enabling continuous improvement of the assessment method
						Requirement	2	8	Public availability
SPICE Conference	2004	A. Anacleto	SPA	Small	4	Observation	1	6	For abroad application in practice, the assessment method requires more detailed guidance
						Observation	1	5	Templates for the documents to be consumed or produced during the assessment
						Requirement	1	7	Assessment output should include process profiles

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Observation	2	4	A mechanism is required which helps to select the key processes based on the business goals and company's problems
						Observation	2	1	Using open style interviews with no questionnaires or checklists because the company representatives usually have low level of software engineering knowledge
						Observation	4	1	More methodological support is required regarding the execution of the interview especially to handle group dynamics
						Observation	1	7	Broader support for the documentation of results of data collection and ratings
						Observation	2	6	Point out the strengths and weakness related to the assessed processes
						Observation	2	6	Point out the risks and improvement suggestions
						Observation	3	1	Systematic tool support for the management of the documents during all assessment activities

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
Info. and Software Tech.	2006	Cater-Steel	SPA	SME	22	Recommendation	4	2	To nurture a relationship of confidence and trust assessors should meet the sponsor prior to the assessment, rather than planning the assessment by phone/email
						Recommendation	4	5	It is suggested that the feedback questionnaire is left with the sponsor at the time of the assessment, and the sponsor return the questionnaire with the comments on the draft of the assessment report
						Recommendation	1	5	Include a template for sponsors to record all costs and benefits from the time of the initial assessment to the follow-up assessment. If firms kept a record of SPI effort, costs and benefits, these details could be summarised in the final report
						Recommendation	1	8	Limited guidance was provided to the follow-up assessors. The program could be improved by including a procedure in the RAPID method for the follow-up meeting

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Recommendation	4	5	Devise two feedback forms for the sponsors: one given to the sponsor at the end of assessment and the other one is given in the follow-up meetings
IATED Conf.	2002	Cater-Steel	SPA	SME	4	Success Factor	5	4	The SPI sponsor in each organization committed the firm to the program
						Success Factor	5	3	Sponsor participated in the planning, assessment and follow-up stages of the program
						Success Factor	4	1	Availability of appropriate training course and seminars for the participants in the organizations
						Success Factor	5	2	Assessors had completed the SPICE certification training
Australian SE Conf,	2004	Cater-Steel	SPA	SME	22	Success Factor	5	4	Commitment of sponsor
						Success Factor	5	3	Participating in the assessment
						Success Factor	4	5	Conduct follow-up meetings to evaluate the degree of improvements achieved

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
SPICE Conf.	2003	Fabbrini	SPA		1	Requirement	2	12	Reliability of SPA method
						Requirement	2	12	Objectiveness and Repeatability of SPA method
						Requirement	2	13	Completeness of SPA method
						Requirement	1	6	Documentability of assessment method
						Requirement	2	10	Cost-effectiveness SPA method
						Requirement	3	1	Assessment Coverage tool
						Requirement	3	1	Usable tool
						Requirement	3	4	Adaptable/Flexible tool
						Requirement	3	1	Meaningful tool able to provide assessor with help
						Requirement	3	1	Inclusive tool that include enough information to increase confidence of taking all aspects into consideration
						Requirement	3	1	Tool that can interact with expert assessors
Software Process Improv. and Practice	1998	K. Kautz	SPI	Small	3	Success Factor	2	4	Flexible, tailored assessment approach
IEEE Software	2007	McCaffery	SPA	Small	2	Demands	2	11	The time to prepare and perform the assessment should be minimal

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Demands	2	4	The assessment method should let companies select process areas that are most relevant to their business goals
CROSS-TALK Journal	2001	Komayama	SPA		1	Requirement	2	4	Customizable Method
						Requirement	2	8	Usable for on-site assessment and self-assessment
						Requirement	2	10	Questions are no more than 150
						Requirement	2	10	Well-structured questionnaire
						Requirement	2	7	A roadmap along with milestones for process improvement should be provided
						Requirement	3	1	Support tools to analyse and visualize the assessment data
						Function	3	1	Assessment data collection and analysis
						Function	3	1	On-line assessment support
						Function	3	1	Analysis and visualization of assessment data
						Function	3	2	Database of historical SPA data
						Function	3	3	Semi-Automatic assessment-report generation
						Requirement	2	10	No many questions

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Function	3	2	The tool build a library of knowledge and experience on process assessment and improvement
						Requirement	4	1	Specific questions should be developed prior to the interview
						Requirement	2	11	Reserve minimum 2 hours for interview
						Requirement	4	4	Order questions based on importance
						Requirement	2	4	Surveillance and interview of general project's status help prioritize questions
						Requirement	2	2	Review documents before the interview
						Requirement	4	1	
						Requirement	4	3	Produce assessment report
						Requirement	5	1	The role or position of the interviewee should be clarified
						Requirement	5	5	A means to indicate the effects of process assessment and improvement quantitatively and objectively should be provided
						Requirement	2	7	It is helpful to prioritize the process improvement actions to be taken

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	5	6	Relationship between product quality or project results and process quality should be clarified based on analysis of project data and assessment results
IEEE Software	1997	Grunbacher	SPA	Small	1	Guide principle	2	10	Assessment of reduced cost and effort
						Guide principle	5	5	Increase awareness and improve knowledge about software quality management.
						Guide principle	5	6	Motivate the staff to further improvement
						Guide principle	2	6	Provide a realistic evaluation of the current status
						Guide principle	2	7	Improvement action plan should be feasible and address the special needs of the small company
						Guide principle	5	3	Participation of top-management, quality management, project management and consultant
						Guide principle	4	1	Hold introductory session
						Guide principle	4	1	Provide training to use the assessment tool
						Guide principle	4	5	Hold follow-up interviews and discussions of the action plan (feedback session)

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						lesson learned	2	11	Self assessment can usually be done in 2 hours
						lesson learned	3	1	Doing an assessment with the help of a tool therefore was highly motivating for the staff
Proceeding of SPICE Conf.	2006	Stanabolian	SPA	Small	46	Strength	2	10	Assessment method is simple and low cost
						Strength	2	12	Assessment method gives accurate insight into the organizations assessed
						Strength	2	10	Simplified vocabulary used making the assessment method understandable
						Strength	2	9	Comply with more formal assessments such as ISO15504 and CMMI
						Strength	4	1	Start by informal assessment method that leads to fully formal assessment method
International conference in software Develop.	2005	Laporte	SPA		23	Lessons Learned	2	10	Simple and low cost assessment method
						Lessons Learned	2	12	Provide accurate insight into assessed organizations
						Lessons Learned	2	4	Context-sensitive assessment approach

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Lessons Learned	2	4	Graduate assessment approach allows different kinds of tailoring to be performed to match available resources
						Lessons Learned	2	10	Simplified vocabulary used
Proceedings of the Conference on product focused SPI	1999	Habra	SPA	SME		Experience factor	2	10	Simple assessment method
Proceedings of EUROSPI'99	1999	Dyba	SPA	SME	2	Lessons Learned	4	1	Involving more than one group in the assessment
						lesson learned	2	4	A team should be put together in the company performing the assessment in order to construct a tailor-made questionnaire
						lesson learned	4	1	
						Lessons Learned	2	10	Time needed to complete the questionnaire should not take more than 30 minutes (around 60 question)
						Lessons Learned	4	1	Hold a presentation for persons participating in the assessment

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Lessons Learned	4	2	There should be a close co-operation and work together between the outsiders (consultants/assessors) and the insiders (organizations participants). (i.e. work to build trust and confidence)
						lesson learned	4	1	During the assessment presentation there will be always discussions about the interpretation of questions, so it is advantageous to conduct the assessment at that time
						Lessons Learned	4	4	All information should be treated as confidential
						Lessons Learned	4	5	Do not wait too long before holding the feedback session
						Lessons Learned	1	6	During the implantation, document any discussions and conclusions of the assessment questions because they will help in the feedback session if the same points are raised again (part of the assessment process documentation)

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Lessons Learned	4	5	During the feedback session, identify the possible areas for improvement. This would help developing successful action plan for improvement
						Lessons Learned	4	5	For the feedback session, choose the suitable number of participants, the group size should be suitable to achieve a successful discussions
IEEE	2006	Ekdahl	SPA	SME	1	lesson learned	4	1	Never underestimate the power of a good plan (planning for the assessment)
						lesson learned	5	1	Local representative in the team facilitates the logistics
						lesson learned	4	2	Open and honest interviews require communication and integrity (trusting relationship)
						lesson learned	3	1	Good tools help
						lesson learned	4	1	Planning and preparing for the interviews by improving the interview questions as our experience improved

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						lesson learned	4	1	Don't miss out on the opportunity to do some training
						lesson learned	1	6	Make sure that the appraisal results are as detailed and actionable as possible. i.e. put more effort into observation forms (documentation) to help organizations know what needs to be done to address a reported weakness)
IEEE	2000	Wieggers et al.	SPA	SME	20	Recommendation	2	12	Develop supporting infrastructure to make the MMA repeatable, reliable and efficient
						Recommendation	4	1	Provide procedural guidance contains detailed procedures for planning and conducting a mini-assessment
						Recommendation	1	6	
						Recommendation	5	2	The procedural guidance documents the roles and activities for the lead and backup assessors
						Recommendation	1	3	

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Recommendation	2	4	Various tables guide the selection of the appropriate component options to satisfy each project's mini-assessment objectives. The tables are updated as new experience is gained and new ways to improve the method are found
						Recommendation	1	6	
						Recommendation	4	1	Develop slide modules that assessors can use during the preparations events
						Recommendation	1	4	Develop slide modules that discuss the confidentiality statement
						Recommendation	4	4	
						Recommendation	3	1	Develop several tools to streamline the execution of a mini-assessment to collect background information prior to the planning meeting
						Recommendation	4	1	Provide the project leader a standard project profile questionnaire and a mini-assessment readiness survey
						Recommendation	1	5	Create forms and templates to plan a mini-assessment, record the assessment time for each assessor and record other summary data

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Recommendation	3	2	Create a database to store information about the conducted mini-assessment
Standard 15504	2003	ISO	SPA			Success Factor	5	4	Commitment
						Success Factor	4	2	Motivation: The organization's management needs to motivate participants to be open and constructive. Process assessments focus on the process, not on the performance employees and hence build bridge of confidence and trust between assessors and assessee
						Success Factor	4	4	Confidentiality
						Success Factor	5	5	Relevance: The Organizational Unit members should believe that the assessment will result in some benefits that will accrue to them directly or indirectly
						Success Factor	5	6	Credibility

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	4	1	Develop assessment plan (includes: assessment input, activities, resources, responsibilities)
						Requirement	1	6	Document assessment plan
						Requirement	1	7	Record assessment data and evidences
						Requirement	2	3	Ensure that the collected data is complete and consistent (data validation and accuracy)
						Requirement	1	7	Record the process attributes rating
						Requirement	4	3	Reporting
						Requirement	5	1	Define organization participants, e.g. sponsor, responsibilities
						Requirement	5	2	Define assessment team responsibilities
						Requirement	4	1	Define assessment input
						Requirement	1	6	Document assessment record
IEEE paper	2001	Cater-Steel	SPA	Small	4	Success Factor	5	4	Sponsor commitment and monitoring

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Success Factor	5	3	Assign responsibilities and sponsor participating in the planning, assessment and follow-up stages of the program
						Success Factor	5	1	
						Success Factor	5	2	Involve experienced and trained assessors
						Success Factor	5	3	Involvement of the company technical staff
IEEE	2005	Choi	SPA/ SPI		3	Recommendation	3	1	Automating the process of process assessment, Visualizing the results of process assessment, Providing various information in software process improvement and Managing the results of process assessment
						Recommendation	3	2	Managing the history of SPI activities through storing the assessment results in a database

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
IEEE	2007	Chen	SPA		1	lesson learned	3	1	Visualized indicators that represent measurement results instantly and allow the user easily identify process improvement opportunities
						lesson learned	3	5	The tool should have well-defined responsibility: user, role and permission management should be well-defined and organized in the system
IEEE	2006	Salvaneschi	SPA/SPI	Small	1	Success Factor	5	4	Management commitment and cooperation of every involved actor
						Success Factor	2	4	Ability to tailor the assessment and improvement to specific goals and needs
IEEE	1997	Lok	SPA		36	Requirement	3	1	Tool Functionality: provide the requisites for the assessment as specified by the requirements
						Requirement	3	5	Support confidentiality of the assessment
						Requirement	3	6	Ensure repeatability of results
						Requirement	3	2	Store data indefinitely without loss
						Requirement	3	1	Tool Usability

Type of Pub	Pub. year	Authors	Context	Org. size	# case studies	Evidence description	ECID	EN	Evidence
						Requirement	3	4	Tool Maintainability: Easy to modify and tailor to specific needs
Elsiver	2008	Pettersson	SPA/SPI	Small	1	Success Factor	5	4	Management Commitment
						Success Factor	5	3	Management Involvement
IEEE	2007	Santos	SPI		>20	Success Factor	4	1	Training the project members for the appraisal interviews that will be carried out during the appraisal
						Success Factor	5	4	The high management commits on all necessary resources

ANNEX VI

THE EVALUATION TOOL DEVELOPED FOR THE EVIDENCE-BASED EVALUATION METHOD

SPA Method Evaluation			
No.	Question	Answer	Comments
SPA-MQ1	Does the method acquire assessment data from interviews?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ2	Does the method acquire assessment data from documents?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ3	Does the method ensure the accuracy of assessment findings	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ4	Is the method flexible and customizable (i.e. possibility of adding new axes) by focusing on high priority processes using certain mechanism?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ5	Does the method provide coverage to a process reference model?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ6	Does the method identify strengths, weaknesses, risks and improvement opportunities?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ7	Does the method Suggest improvement action plan?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ8	Is the method publically available and usable for in-site and self assessments?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ9	Does the method comply with formal assessment method	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ10	Is the method simple and well structured with no more than 150 questions in the questionnaire?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ11	Does the assessment duration takes between 2 – 8 hours?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ12	Does the assessment method ensure, based on studies, Reliability?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-MQ13	Does the method ensure completeness?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	

SPA Tool Evaluation			
No.	Question	Answer	Comments
SPA-TQ1	Does the tool cover different assessment phases including collect, analyze and visualize assessment data?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-TQ2	Does the tool create a database of historical SPA data?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-TQ3	Does the tool generate a semi-automatic assessment report?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-TQ4	Is the tool adaptable / Flexible (i.e. by adding new axes)?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-TQ5	Does the tool support the confidentiality of assessment	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-TQ6	Does the method ensure the repeatability of results	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA Procedure Evaluation			
No.	Question	Answer	Comments
SPA-PQ1	Does the procedure allow preparing participants in the assessment process?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-PQ2	Does the SPA procedure work to build a relationship of trust and confidence between the assessors and sponsors prior to the assessment?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-PQ3	Does the procedure produce assessment report and deliver it to the organization?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-PQ4	Does the assessment procedure ensure the confidentiality of the participants (even before conducting the assessment)	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-PQ5	Does the assessor hold feedback session after assessment to keep improving the assessment method	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA Documentation Evaluation			
No.	Question	Answer	Comments
SPA-DQ1	Is the assessment purpose, objectives and needed resources are documented?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-DQ2	Is the guidance for identifying assessed organization unit provided (the assessment scope)?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	

SPA Documentation Evaluation			
No.	Question	Answer	Comments
SPA-DQ3	Is the guidance for the assessment team provided?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-DQ4	Is the confidentiality of the assessment documented?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-DQ5	Does the method provide all necessary documents and templates?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-DQ6	Is the assessment process documented?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-DQ7	Are the data collection results and ratings documented for each organization?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-DQ8	Is the guidance for the follow-up meetings provided	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA User Evaluation			
No.	Question	Answer	Comments
SPA-UQ1	Are the assessment sponsor and participants (interviewee) responsibilities defined?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-UQ2	Are the assessment team credentials and responsibilities defined?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-UQ3	Is the senior management and other staff members involved in the assessment process?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-UQ4	Does the participants commitment ensured	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-UQ5	Do the sponsor and participants believe that the assessment will give results?	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	
SPA-UQ6	Does the assessment participant feel the benefits of conducting an assessment	<input type="radio"/> F <input type="radio"/> P <input type="radio"/> N	

BIBLIOGRAPHIE

- Abrahamson, P., O. Salo, et al. (2002). Agile Software Development Methods, Review and Analysis. VTT Publications 478. Espo, Finland: 107 p.
- Abran, A. and A. April (2006). Assessment of Maintenance using a Software Maintenance Maturity Model (S3m) IT Productivity Metrics, Toronto, Canada.
- Abran, A. and K. Meridji (2006). Analysis of Software Engineering from An Engineering Perspective. UPGRADE. VII. p. 5-19
- Abran, A., J. Moore, et al. (2004). Guide to the Software Engineering Body of Knowledge – SWEBOK, 2004 Version, IEEE Computer Society Press, URL: <http://www.swebok.org>.
- ACUÑA, S. T., A. D. ANTONIO, et al. (2000). "THE SOFTWARE PROCESS: MODELLING, EVALUATION AND IMPROVEMENT." Handbook of Software Engineering and Knowledge Engineering 0(0).
- Ahren, D. M., A. Clouse, et al. (2004). CMMI Distilled, A practical introduction to integrated process improvement, Addison Wesley - book's web site.
- Alexandre, S., A. Renault, et al. (2006). "OWPL: A Gradual Approach for Software Process Improvement In SMEs." Proceedings of the 32nd EUROMICRO Conference on Software Engineering and Advanced Applications (EUROMICRO-SEAA'06): p. 328-335.
- Anacleto, A., C. Wangenheim, et al. (2004). A method for Process Assessment in Small Software Companies. 4th International SPICE Conference on Process Assessment and Improvement, Portugal.
- Anacleto, A., C. G. Wangenheim, et al. (2004). Experiences gained from applying ISO/IEC 15504 to small software companies in Brazil. 4th International SPICE Conference on Process Assessment and Improvement. Lisbon, Portugal.
- April, A. (2005). SMcmm: Software Maintenance Capability Maturity Model. Statistic Canada Conference, Ottawa, Canada.
- April, A. and A. Abran (2008). Software Maintenance Management: Evaluation and Continuous Improvement, John Wiley & Sons, Inc., Hoboken, New Jersey.
- April, A., A. Abran, et al. (2004). Assessment of Software Maintenance Capability: A model and its Architecture. IASTED conference on Software Engineering, Innsbruck, Austria.

- April, A., A. Abran, et al. (2004). SMmmm Model to Evaluate and Improve the Quality of the Software Maintenance Process. 8th European Conference on Software Maintenance and Re-Engineering. Tampere, Finland: p. 243-248.
- April, A., A. Abran, et al. (2004). Software Maintenance Productivity measurement: how to Assess the readiness of your organization. 14th International Workshop on Software Measurement (IWSM) IWSM-Metrikon 2004, Konigs Wusterhausen, Magdeburg, Germany, Shaker-Verlag.
- April, A. and J.-M. Desharnais (2005). SOFTWARE MAINTENANCE MATURITY MODEL (SMmmm): A SOFTWARE MAINTENANCE PROCESS MODEL. Statistic Canada Conference, Ottawa, Canada.
- April, A., J.-M. Desharnais, et al. (2006). A Formalism of ontology to support a software maintenance knowledge-based system. Proceedings of the Eighteenth International Conference on Software Engineering & Knowledge Engineering Conference (SEKE06), San Francisco, CA, USA.
- April, A., J. H. Hayes, et al. (2005). "Software Maintenance Maturity Model (SMmmm): the software maintenance process model." Journal of Software Maintenance and Evolution: Research and Practice 17(3): p. 197 - 223.
- April, A. H., J. Huffman, et al. (2005). "Software Maintenance Maturity Model (SMmmm): The software maintenance process model." Journal of Software Maintenance and Evolution: Research and Practice 17(3): p. 197-223.
- Ares, J., R. Garcia, et al. (2000). "A more rigorous and comprehensive approach to software process assessment." Software Process - Improvement and Practice 5(1): p. 3-30.
- Ashrafi, N. (2003). "The impact of software process improvement on quality: in theory and practice." Information & Management 40: p. 677-690.
- Basili, V. (1985). Quantitative Evaluation of Software Engineering Methodology. The First Pan Pacific Computer Conference Melbourne, Australia.
- Basili, V. R. and G. Caldiera (1995). Improve Software Quality by Reusing Knowledge and Experience. Sloan Management Review/Fall: p. 55-64.
- Basili, V. R. and A. Turner (1975). "Iterative enhancement: A Practical Technique for Software Engineering." IEEE Transactions on Software Engineering 1(4): p. 110-125.
- Beitz, A., K. E. Emam, et al. (1999). A business focus to assessments. Proceedings of the European Conference on Software Process Improvement.
- Bhandari, I., M. Halliday, et al. (1993). "A case study of software process improvement during development." Software Engineering, IEEE Transactions 19(12): p. 1157 - 1170.

- Boehm, B. W. (1988). "A Spiral Model for Software Development and Enhancement " IEEE Computer 21(5): p. 61-72.
- Bollinger, T. B. and C. McGowan (1991). "A critical look at software capability evaluations." IEEE Software 8(4): p. 25 - 41.
- Bryant, A. (2002). "Grounding Systems Research: Re-establishing Grounded Theory." Proceedings of the 35th Hawaii International Conference on System Sciences: p. 253-263.
- Card, D. N. (2004). Research Directions in Software Process Improvement. Proceedings of the 28th Annual International Computer Software and Applications Conference (COMPSAC'04), Hong Kong, IEEE.
- Carver, J. (2007). The Use of Grounded Theory in Empirical Software Engineering Empirical Software Engineering Issues. Critical Assessment and Future Directions, Springer Berlin / Heidelberg.
- Cass, A., C. Volcker, et al. (2000). SPiCE for SPACE: A method of process assessment for space software projects, SPICE 2000, Ireland.
- Cater-Steel, A. (2002). Process capability assessments in small development firms. IASTED 6th International Conference Software Engineering and Applications, Cambridge, MA, USA.
- Cater-Steel, A., M. Toleman, et al. (2006). "Process improvement for small firms: An evaluation of the RAPID assessment-based method." Elsevier, Information and Software Technology 48: p. 323–334.
- Cater-Steel, A. P. (2001). Process Improvement in Four Small Software Companies. 13th Australian Software Engineering Conference (ASWEC'01). Canberra, Australia, IEEE.
- Cater-Steel, A. P. (2004). An Evaluation of Software Development Practice and Assessment-Based Process Improvement in Small Software Development Firms. School of Computing and Information Technology/ Faculty of Engineering and Information Technology. Queensland - Australia, Griffith University.
- Cater-Steel, A. P. (2004). Low-rigour, Rapid Software Process Assessments for Small Software Development Firms. Proceedings of the 2004 Australian Software Engineering Conference (ASWEC'04). Melbourne, Australia, IEEE Computer Society.
- Cepeda, S., M. J. Staley, et al. (2004). SE2 Workshop on applying CMMI in small business settings, CMU/SEI.

- CETIC (2006). OWPL Software Process Improvement for VSE, SME and low maturity enterprises Version 1.2.2. Namur, Belgium, University of Namur, Software Quality Lab: 50 p.
- Chen, X., P. Sorenson, et al. (2007). Continuous SPA: Continuous Assessing and Monitoring Software Process. 2007 IEEE Congress on Services (Services 2007), IEEE: p. 153-158.
- Choi, Y., E. Lee, et al. (2005). The Management of Software Processes with Software Process Improvement Tool based on ISO 15504. The 7th International Conference on Advanced Communication Technology, ICACT 2005, Korea, IEEE.
- Cignoni, G. A. (1999). Rapid Software Process Assessment to promote Innovation in SMEs. EUROMICRO'99, Milan, Italy, 8-10, September 1999.
- CMMI-SUT (2006). CMU/SEI-2006-TR-011: Appraisal Requirements for CMMI, Version 1.2 (ARC, V1.2), SCAMPI Upgrade Team. Pittsburgh, PA, Carnegie Mellon Software Engineering Institute.
- CMMI Product Team (2006). CMMI for Development Version 1.2. Pittsburgh, PA, Carnegie Mellon, Software Engineering Institute.
- CMU/SEI (2006). CMU/SEI-2006-TR-011: Appraisal Requirements for CMMI, Version 1.2 (ARC, V1.2), SCAMPI Upgrade Team. Pittsburgh, PA, Carnegie Mellon Software Engineering Institute.
- Curtis, B., H. Krasner, et al. (1987). On Building Software Process Models under the Lamppost. Proceedings of the 9th International Conference on Software Engineering (ICSE 1987). Monterey, CA, IEEE Computer Society Press: p. 96-103.
- Curtis, W. (1994). At the Plenary Session: setting an agenda for measurement research and practice, at the 2nd International Symposium on Software Metrics: Bridging the Gap Between Research and Practice, London, UK.
- Dangle, K. C., P. Larsen, et al. (2005). "Software process improvement in small organizations: a case study." IEEE Software 22(6): p. 68-75.
- Diaz, M. and J. Sligo (1997). "How software process improvement helped Motorola." Software, IEEE 14(5): p. 75-81.
- Dyba, T. (2000). "An Instrument for Measuring the key Factors of Success in software Process Improvement." Empirical Software Engineering 5: p. 357-390.
- Dyba, T. and N. B. Moe (1999). Rethinking the Concept of Software Process Assessment. Proceedings of European Software Process Improvement Conference EuroSPI '99.

- Ekdahl, F. and S. Larsson (2006). Experience Report: Using Internal CMMI Appraisals to Institutionalize Software Development Performance Improvement. Proceedings of the 32nd EUROMICRO Conference on Software Engineering and Advanced Applications (EUROMICRO-SEAA'06), Cavtat/Dubrovnik, Croatia, IEEE.
- El-Emam, K., L. Briand, et al. (1996). "Assessor agreement in rating SPICE processes." Software Process - Improvement and Practices 2 (4): p. 291-306.
- El-Emam, K., J.-N. Drouin, et al. (1998). SPICE - The Theory and Practice of Software Process Improvement and Capability Determination. Los Alamitos CA, Wiley-IEEE Computer Society Press.
- El-Emam, K. and D. R. Goldenson (1995). SPICE: An Empiricist's Perspective. Proceedings of the Second International Software Engineering Standards Symposium, Montreal, Canada.
- El-Emam, K. and D. R. Goldenson (1999). An Empirical Review of Software Process Assessments, National Research Council of Canada, Institute for Information Technology: 86 p.
- El-Emam, K. and N. H. Madhavji (1999). The reliability of measuring organizational maturity. Software Process Improvement and Practice, John Wiley & Sons.
- Fabbrini, F., E. Fantini, et al. (2003). Performing SPICE Assessments: Yet Another Tool. In Joint ESA - 3rd International SPICE Conference on Process Assessment and Improvement, ESTEC, Noordwijk, The Netherlands.
- Fuggetta, A. (2000). Software Process: A Roadmap. Proceedings of the conference on The future of Software engineering, Limerick, Ireland, New York: ACM Press.
- FUNDP-CETIC – Software Quality Lab (2006). OWPL Software Process Improvement for VSE, SME and low maturity enterprises Version 1.2.2, University of Namur.
- FUNDP-CETIC (2006). OWPL Software Process Improvement for VSE, SME and low maturity enterprises.
- Gilb, T. (1988). Principles of Software Engineering Management, Addison-Wesley. Reading, MA.
- Gilgun, J. F. (1992). Definitions, Methodologies, and Methods in Qualitative Family Research. Qualitative Methods in Family Research, Thousand Oaks: Sage.
- Glaser, B. (1978). Theoretical Sensitivity: Advances in the methodology of Grounded Theory, Sociology Press.
- Glasser, B. and A. Strauss (1965). Awareness of Dying, Aldine.

- Glasser, B. and A. Strauss (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research, Aldine.
- GMOD (1992). V-Model: Software Lifecycle Process Model, General Report No. 250, German Ministry of Defence.
- Goldenson, D. R., K. E. Emam, et al. (1997). Empirical Studies of Software Process Assessment Methods, International Software Engineering Research Network.
- Goldenson, D. R. and J. D. Herbsleb (1995). After the Appraisal: A Systematic Survey of Process Improvement, its Benefits, and Factors that Influence Success. CMU/SEI-95-TR-009. Pittsburgh, Pennsylvania 15213, Software Engineering Institute/Carnegie Mellon University.
- Gorschek, T. (2006). Requirements Engineering Supporting Technical Product Management. Department of Systems and Software Engineering. Karlskrona, Sweden, Blekinge Institute of Technology: 343 p.
- Gray, E. M. and W. L. Smith (1998). "On the limitations of the software process assessment and the recognition of a required re-orientation for global process improvement." Software Quality Journal 7: p. 21-34.
- Grünbacher, P. (1997). A software assessment process for small software enterprises. 23rd EUROMICRO Conference '97, IEEE Computer Society.
- Habra, N., Eustache Niyitugabira, et al. (1999). Software Process Improvement in Small Organizations Using Gradual Evaluation Schema. Proceedings of the International Conference on Product Focused Software Process Improvement, Oulu, Finland.
- Habra, N., E. Niyitugabira, et al. (1999). Software Process Improvement for Small Structures: First Results of a Micro-Assessment Framework. Proceedings of the European Conference on Software Process Improvement SPI'99. Barcelona, Spain.
- Habra, N. and A. Renault (2004). OWPL Une Méthodologie et des Modèles Légers pour Initier une Démarche d'Amélioration des Pratiques Logicielles APL. Ingénierie des Systèmes d'Information.
- Habra, N., A. Renault, et al. (2002). OWPL Micro Assessment. Proceedings for Software Quality Workshop, 24rd International conference on Software Engineering ICSE, Orlando, Florida USA, ACM.
- Halvorsen, C. P. and C. Reidar (1999). A Taxonomy for SPI Frameworks. 24th NASA Software Engineering Workshop, Greenbelt/Washington, USA.
- Hamann, D., A. Beitz, et al. (2001). Using FAME Assessments to Define Measurement Goals. International Workshop on Software Measurement. R. Dumke and A. Abran. Berlin, Germany, Springer-Verlag Berlin Heidelberg: p. 220-232.

- Herbsleb, J., A. Carleton, et al. (1994). Benefits of CMM-Based Software Process Improvement: Initial Results, Technical Report: CMU/SEI-94-TR-013, CMU/SEI.
- House, E. (1980). Evaluating With Validity, Sage Pubns.
- Humphrey, W. S. and D. Kitson (1987). CMU/SEI-87-TR-16, ESD-TR-87-117: Preliminary Report on Conducting SEI-assisted Assessments of software Engineering Capability.
- Humphrey, W. S. (1989). Managing the software Process. In: SEI Series in Software Engineering. Massachusetts, Addison Wesley Publishing Company.
- Humphrey, W. S. (1992). CMU/SEI-92-TR-7: Introduction to Software Process Improvement. CMU/SEI-TR-92-7. Pittsburgh, Pennsylvania, CMU/SEI.
- Humphrey, W. S. and D. Kitson (1987). CMU/SEI-87-TR-16, ESD-TR-87-117: Preliminary Report on Conducting SEI-assisted Assessments of software Engineering Capability.
- ISO-9001 (1993). Quality systems - Model for quality assurance in design, development, production, installation and servicing, International Organization for Standardization.
- ISO/IEC (1992). ISO/IEC JTC1/SC7 N944R, The Need and Requirements for a Software Process Assessment Standard. Study Report, Issue 2.0, June 1992.
- ISO/IEC (1998). "ISO/IEC TR 15504, Information Technology - Software Process Assessment - Parts 1-9".
- ISO/IEC (2003). ISO/IEC 15504-2, Software Engineering - Process Assessment - Part 2: Performing An Assessment.
- ISO/IEC (2003). ISO/IEC 15504-3 Information technology — Process Assessment — Part 3: Guidance on performing an assessment.
- ISO/IEC (2003-2006). ISO/IEC 15504 Information Technology - Process Assessment - Parts 1-5.
- ISO/IEC (2004). ISO/IEC 15504-1 Information Technology — Process Assessment — Part 1: Concepts and Vocabulary.
- Johnson, D. L. and J. G. Brodman (1997). Tailoring the CMM for Small Businesses, Small Organizations, and Small Projects. Software Process Newsletter, No. 8, Winter 1997, IEEE Computer Society.
- Jorgensen, N. (2005). The Engineering of Software: Views from Technology Theory. Twenty-Sixth International Conference on Information Systems, Las Vegas, USA.

- Jung, H.-W. and D. R. Goldenson (2002). The Internal Consistency of Key Process Areas in the Capability Maturity Model (CMM) for Software (SW-CMM), Technical Report: CMU/SEI-2002-TR-037, CMU/SEI.
- Kautz, K. (1998). "Software Process Improvement In Very Small Enterprises: Does It Pay Off?" SOFTWARE PROCESS—Improvement and Practice 4(4): p. 209-226.
- Kautz, K. and F. Ramzan (2001). Software Quality Management and Software Process Improvement in Denmark. Proceedings of the 34th Hawaii International Conference on System Sciences, Maui, Hawaii, USA, IEEE Computer Society.
- Kelly, D. P. and B. Culleton (1999). "Process Improvement for Small Organization." Computer 32(10): p. 41-47.
- Kim, H. K. and H. S. Yang (2006). Software Process Improvement Environment. ICCSA 2006, Glasgow, UK Springer-Verlag Berlin Heidelberg 2006.
- Kitchenham, B. (2007). Guidelines for Performing Systematic Literature Reviews in Software Engineering, Joint Technical Report, Keele University, EBSE-2007-01, July 2007.
- Knauber, P., D. Muthig, et al. (2000). "Applying Product Line Concepts in Small and Medium-Sized Companies." IEEE Software: p. 88-95.
- Komi-Servio, S. (2004). Development and Evaluation of Software Process Improvement Methods. VTT Technical Research Center with Assent of the Faculty of Science. Linnanmaa, Finland, University of Oulu: 190 p.
- Komiyama, T., T. Sunazuka, et al. (2001). "Proposal on Library-Centered Software Process Assessment." CROSSTALK The Journal of Defense Software Engineering 14(8): p. 22-28.
- Kuilboer, J. P. and N. Ashrafi (2000). "Software process and product improvement: an empirical assessment." Information and Software Technology 42(1): p. 27-34.
- Kuvaja, P., J. Palo, et al. (1999). "TAPISTRY - A Software Process Improvement Approach Tailored for Small Enterprises." Software Quality Journal 8(2): p. 149-156.
- Laporte, C. and A. April (2005). Applying Software Engineering Standards in Small Settings: Recent historical perspectives initial achievements. International Research Workshop for Process Improvement in Small Settings, Software Engineering Institute, Pittsburgh.
- Laporte, C. Y., J.-M. Desharnais, et al. (2005). Initiating Software Process Improvement in Small Enterprises: Experiments with Micro-Evaluation Framework. Proceedings of the SWDC-REK International Conference on Software Development, Reykjavik, Iceland.

- Laryd, A. and T. Orci (2000). Dynamic CMM for Small Organizations. Proceedings of the First Argentine Symposium on Software Engineering - ASSE 2000, Tandil, Argentina.
- Lee, H.-y., H.-W. Jung, et al. (2001). Analysis of interrater agreement in ISO/IEC 15504-based software process assessment. Second Asia-Pacific Conference on Quality Software (APAQS'01) HONG KONG IEEE.
- Lok, R. H. and A. Walker (1997). Automated Tool Support for an Emerging International Software Process Assessment Standard. Proceedings of the 3rd International Software Engineering Standards Symposium (ISESS '97), Walnut Creek, California, USA, IEEE Computer Society.
- Loon, H. v., A. Cass, et al. (2004). Process Assessment and Improvement: A practice guide for managers, quality professional and assessors. NY, USA, Springer.
- López, M. (2000). An Evaluation Theory Perspective of the Architecture Tradeoff Analysis Method (ATAM). CMU/SEI-2000-TR-012. Pittsburgh, PA, CMU/SEI.
- López, M. (2003). "Application of an evaluation framework for analyzing the architecture tradeoff analysis method." The Journal of Systems and Software 68: p. 233-241.
- M. C. Paulk et al. (1995). The Capability Maturity Model: Guidelines for Improving the Software Process, Addison Wesley, Reading, Mass.
- Mäkinen, T., T. Varkoi, et al. (2000). A Detailed Process Assessment Method For Software SMEs European Software Process Improvement SPI and Assessments (EuroSPI 2000). Copenhagen, Denmark.
- McCaffery, F., D. McFall, et al. (2005). "Improving the Express Process Appraisal Method." PROFES: p. 286-298.
- McCaffery, F., P. Taylor, et al. (2007). "Adept: A Unified Assessment Method for Small Software Companies." IEEE Software Special Issue of "SE Challenges in Small Software Organizations": p. 24-31.
- McGarry, F., R. Pajerski, et al. (1994). Software Process Improvement in the NASA Software Engineering Laboratory, Technical Report: CMU/SEI-94-TR-22, CMU/SEI: 80 p.
- Mills, H. D., D. O'Neill, et al. (1980). "The Management of Software Engineering." IBM System Journal 24(2): p. 414-477.
- Morse, J., J. Swanson, et al. (2001). The Nature of Qualitative Evidence. Thousand Oaks, California, Sage Publications, Inc.

- Morse, J. M., M. Barrett, et al. (2002). "Verification Strategies for Establishing Reliability and Validity in Qualitative Research." International Journal of Qualitative Methods 1(2).
- Niazi, M., D. Wilson, et al. (2003). A model for the implementation of a software process improvement: A pilot study. Proceedings of the Third International Conference on Quality Software (QSIC'03), Dallas, Texas, USA, IEEE Computer Society, Washington, DC, USA.
- Niazi, M., D. Wilson, et al. (2005). "A framework for assisting the design of effective software process improvement implementation strategies." The Journal of Systems and Software 78(2): p. 204-222.
- Niazi, M., D. Wilson, et al. (2005). "A maturity model for the implementation of software process improvement: an empirical study " Journal of Systems and Software 74(2): p. 155-172.
- Niazi, M., D. Wilson, et al. (2006). "Critical Success Factors for Software Process Improvement Implementation: An Empirical Study." SOFTWARE PROCESS IMPROVEMENT AND PRACTICE 11(2).
- Paquette, D.-A., A. April, et al. (2006). "Assessment Results using the Software Maintenance Maturity Model (S3m)." In the Preceedings of the 16th International Workshop on Software Measurement - (IWSM-Metrikom 2006): p. 147-160.
- Paulk, M. C. (1995). "How ISO 9001 Compares with the CMM." IEEE Software 12(1): p. 74-82.
- Pettersson, F., M. Ivarsson, et al. (2008). "A practitioner's guide to light weight software process assessment and improvement planning." The Journal of Systems and Software 81: p. 972-995.
- Pikkarainen, M. (2006). Agile Assessment Approach (based on the eight case experiences), Technical Report: D4.2_v1.0, Technical Research Centre of Finland, VTT: 20 p.
- Rainer, A. and T. Hall (2003). "A quantitative and qualitative analysis of factors affecting software processes." The Journal of Systems and Software 66(1): p. 7-21.
- Renault, A. (1999). Micro-évaluation des pratiques logicielles - Etude de cas Institute d'Informatique FUNDEP - Namur: 7 p.
- Richardson, I. (2001). "Software Process Matrix: A Small Company SPI Model." Software Process Improvement Practices 6: p. 157-165.
- Richardson, I. (2002). "SPI Models: What Characteristics are Required for Small Software Development Companies?" Software Quality Journal 10: p. 101-114.

- Rout, T. P., A. Tuffley, et al. (2000). The Rapid Assessment of Software Process Capability. First International SPICE Conference, Limerick, Ireland.
- Rozman, I., R. V. Horvat, et al. (1994). "UNITED VIEW ON ISO 9001 MODEL AND SEI CMM." IEEE: p. 56-63.
- Salvaneschi, P., D. Grasso, et al. (2006). "The Role of Process Improvement in IT Departments: Experiences and Lessons Learned." Softw. Process Improve. Pract. 11(3): p. 309–317.
- Santos, G., M. Montoni, et al. (2007). SPI-KM - Lessons Learned from Applying a Software Process Improvement Strategy Supported by Knowledge Management. PROFES 2007, Riga, Latvia, Springer-Verlag Berlin Heidelberg 2007.
- Santos, G., M. Montoni, et al. (2007). Implementing Software Process Improvement Initiatives in Small and Medium-Size Enterprises in Brazil. 6th International Conference on the Quality of Information and Communications Technology, 2007. QUATIC 2007. , Lisbon, Portugal, IEEE.
- Scriven, M. (1991). Evaluation Thesaurus. Newbury Park, CA, Sage Publications.
- Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary vision. In S. I. Evaluating social programs and problems: Visions for the new millennium Donaldson & M. Scriven (Eds.). Mahwah, NJ: Lawrence Erlbaum Associates.: p. 19-42.
- Seaman, C. (1999). "Qualitative Methods in Empirical Studies of Software engineering." IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 25(4): p. 557-572.
- Shadish, W., T. Cook, et al. (1993). Foundations of Program Evaluation. Theories of Practice. Newbury Park, CA, Sage Publications.
- Shaw, M. (1990). "Prospects for an Engineering Discipline of Software." IEEE Software 7 no. 6: p. 15-24.
- Shaw, M. (2002). "What makes good research in software engineering?" International Journal of Software Tools for Technology Transfer 4(1): p. 1-7.
- Simon, H. (1996). The Sciences of the Artificial. Cambridge, MA, MIT Press.
- Sørungård, S. (1997). Verification of Process Conformance in Empirical Studies of Software Development. Department of Computer and Information Science, The Norwegian University of Science and Technology: 252 p.
- SPICE. (1998). "Software Process Improvement and Capability dEtermination Web Site [Online]." Retrieved July 10, 2008, from: <http://www.sqi.gu.edu.au/spice/suite/intro.html>.

- Stambollian, A., N. Habra, et al. (2006). OWPL: A Light Model & Methodology for Initiation Software Process Improvement. Proceedings of SPICE Conference, Luxembourg.
- Staples, M. and M. Niazi (2007). "Experience using systematic review guidelines." The Journal of Systems and Software, Elsevier Science Inc. 80(9): p. 1425-1437.
- Taylor-Powell, E. and S. Steel (1996). Program Development and Evaluation, Collecting Evaluation Data: An Overview of Sources and Methods, Report: G3658-4, Available at <http://learningstore.uwex.edu/pdf/G3658-4.pdf>, University of Wisconsin-Extension.
- Thomas, M. and F. McGarry (1994). "Top-Down vs. Bottom-Up Process Improvement." IEEE Software 11(4): p. 12-13.
- Tingey, M. O. (1997). Comparing ISO 9000, Malcolm Baldrige, and the SEI CMM for software: a reference and selection guide, Upper Saddle River: Prentice-Hall, Inc.
- TOMASO, C. D. (2008). Analyse de contexte et évolution d'un outil expérimental de support aux évaluations de la maturité des processus de la maintenance : S3mAssess. Institut d'Informatique. Namur, Belgium, Facultés Universitaires Notre-Dame de la Paix 325 p.
- Varkoi, T. and T. Makinen (2000). Software process improvement initiation in small organisations. The 3rd European Software Measurement Conference, FESMA-AEMES Madrid, Spain.
- Villalón, J. A., A. G. Cuevas, et al. (2002). "Experiences in the Application of Software Process Improvement in SMES." Software Quality Journal 10(3): p. 261-273.
- Vincent, L. (2008). Méthode d'évaluation des processus de la maintenance du logiciel. Institut d'Informatique. Namur, Belgium, Facultés Universitaires Notre-Dame de la Paix 224 p.
- Vincenti, W. G. (1990). What engineers know and how they know it. Baltimore, London, The John Hopkins University Press.
- Volcker, C. and A. Cass (2000). PASCON/W06 CCN5/TN7: ISO/IEC TR 15504 Conformant Method for the Assessment of Space Software Processes. Technical Note No. 7. Part A: SPiCE for SPACE Assessment Model and Part B: SPiCE for SPACE Method, Issue 1.1, December 19, 2002.
- Volcker, C. and A. Cass (2001). PASCON/W06-CCN5/TN8: Guidelines for Space Software Process Improvement, Technical Note 8, Issue 1.0, December 19, 2001.

- Wangenheim, C., A. Anacleto, et al. (2004). MARES - A Methodology for Software Process Assessment in Small Software Companies, Technical Report: LQPS001_04E. Brazil, LPQS - Universidade do Vale do Itajai.
- Wangenheim, C., T. Varkoi, et al. (2006). "Standard based software process assessments in small companies." SOFTWARE PROCESS IMPROVEMENT AND PRACTICE 11(3): p. 329-335.
- Wangenheim, C., S. Weber, et al. (2006). "Experiences on establishing software processes in small companies " Information and Software Technology 48(9): p. 890-900.
- Wangenheim, C. G., A. Anacleto, et al. (2006). "Helping small companies assess software processes." IEEE Software 23(1): p. 91-98.
- Wangenheim, C. G. v., T. Varkoi, et al. (2005). Performing ISO 15504 Conformant Software Process Assessment in Small Software Companies. Proceedings EUROSPI'2005 Conference, Hungary.
- Ward, R. P., M. Laitinen, et al. (2000). "Thinking Objectively - Management in the Small." COMMUNICATIONS OF THE ACM 43: p. 113-116.
- Webster, J. and R. T. Watson (2002). "Analyzing the Past to Prepare for the Future: Writing a Literature Review." MIS Quarterly 22(2): p. xiii-xxiii.
- Wieggers, K. E. and D. C. Sturzenberger (2000). "A modular Software Process Mini-Assessment Method." IEEE Software 17(1): p. 62-69.
- Wilkie, F. G., F. McCaffery, et al. (2007). "A Low-overhead Method for Software Process Appraisal." Software Process - Improvement and Practice 12(4): p. 339-349.
- Worthen, B., J. Sanders, et al. (1997). Program Evaluation. Alternative Approaches and Practical Guidelines. New York, N.Y., Addison Wesley Longman.
- Zahran, S. (1998). Software Process Improvement: Practical Guidelines for Business Success, . Harlow, England, Addison-Wesley.