An Empirical Investigation of the Key Factors for Success in Software Process Improvement

Tore Dybå, *Member*, *IEEE Computer Society*

Abstract—Understanding how to implement software process improvement (SPI) successfully is arguably the most challenging issue facing the SPI field today. The SPI literature contains many case studies of successful companies and descriptions of their SPI programs. However, the research efforts to date are limited and inconclusive and without adequate theoretical and psychometric justification. This paper extends and integrates models from prior research by performing an empirical investigation of the key factors for success in SPI. A quantitative survey of 120 software organizations was designed to test the conceptual model and hypotheses of the study. The results indicate that success depends critically on six organizational factors, which explained more than 50 percent of the variance in the outcome variable. The main contribution of the paper is to increase the understanding of the influence of organizational issues by empirically showing that they are at least as important as technology for succeeding with SPI and, thus, to provide researchers and practitioners with important new insights regarding the critical factors of success in SPI.

Index Terms—Empirical software engineering, software process improvement, critical success factors, organizational issues, survey research.

1 Introduction

During the last decade, the software industry has been more and more concerned about software process improvement (SPI). Consequently, we have witnessed a proliferation of models and initiatives all claiming to increase the likelihood of succeeding with SPI initiatives, e.g., the Quality Improvement Paradigm (QIP) [7], the IDEAL model [70], the *ami* method [78], and SPICE [42]. The SPI literature is full of case studies and anecdotal evidence of successful companies and descriptions of their SPI programs, e.g., Alcatel [30], Hewlett-Packard [48], Hughes [57], Motorola [27], NASA [9], Philips [81], Raytheon [34], and Siemens [71]. Several authors repeatedly discuss the importance of certain success factors (e.g., [31], [43], [44], [47], [48], [79], [86], [87]. However, the research efforts to date are limited and inconclusive and without adequate theoretical and psychometric justification [35]. Even for commonly recognized factors such as management commitment and employee participation, no operational measures are available [2].

SPI has its roots in quality management and is closely related to "second generation" [46] organizational development approaches, specifically to organizational learning. Understanding the "quality revolution" is, therefore, an important prerequisite for understanding SPI.

The current state-of-the-art in quality management has more than anything else been shaped by quality gurus such as Deming [33], Juran [58], Crosby [25], and their quality frameworks. These and other authors (e.g., [3], [15], [77], [82], [97]) repeatedly discuss the importance of critical factors such as leadership involvement, employee participation, measurement, and process management to improve the quality performance in organizations.

Similarly, the current state-of-the-art in organizational learning has more than anything else been shaped by the works of Argyris and Schön, who made an important distinction between the concepts of *single-loop* and *double-loop* learning [5], Senge, who proposed five disciplines for creating a learning organization [84], and Nonaka, who discussed the conditions required of the "knowledge creating company" [74].

Humphrey [55], [56] and Basili [7], [8] have been the pioneers and leaders in the field of SPI, identifying the basic principles of software process change and improvement, goal-oriented measurement, and the reuse of organizational experience and learning. For a more detailed review of these and other studies of the facilitating factors of quality management, organizational learning, and SPI, see [35].

This paper extends and integrates models from prior research by performing an empirical investigation of the key factors for success in SPI. In the next section, we present our proposed model, which comprises six independent, one dependent, and two contextual variables, and the corresponding research hypotheses. In Section 3, we provide an overview of the population and sample, the variables and measures, the data collection procedure, the reliability and validity of the measurement scales, and the data analysis techniques. Section 4 presents the results of testing the hypothesis and exploring the relationships. Section 5 provides a discussion of the results, their implications, the limitations of the study, and suggests some directions for further research. Section 6 provides some concluding comments.

- *The author is with the Software Engineering Department, SINTEF ICT, NO-7465 Trondheim, Norway. E-mail: tore.dyba@sintef.no.*

Manuscript received 27 Aug. 2004; revised 15 Feb. 2005; accepted 22 Apr. 2005; published online 26 May 2005.
Recommended for acceptance by P. Jalote.
For information on obtaining reprints of this article, please send e-mail to: tse@computer.org, and reference IEEECS Log Number TSE-0179-0804.

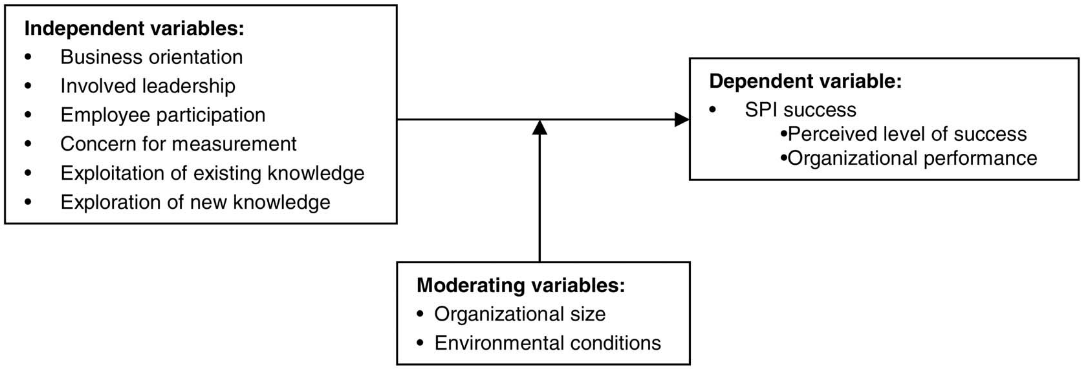

Fig. 1. Conceptual research model.

2 Conceptual Model and Hypotheses

The research model to be empirically tested in this study is depicted in Fig. 1. The model derives its theoretical foundations by combining prior research in quality management, organizational learning, and SPI with an explorative study of 54 managers, developers, and customer representatives in four software companies along with a detailed review by eleven SPI experts from academia and industry (see [35] and [37]).

The model includes six facilitating factors: business orientation, involved leadership, employee participation, concern for measurement, exploitation of existing knowledge, and exploration of new knowledge, two moderating factors (organizational size and environmental conditions), and the dependent variable, SPI success. These are described next.

2.1 Business Orientation

Attaining business objectives is the ultimate target for any change program [10], and the role of process orientation and process improvement has long been recognized as essential factors to achieve business excellence and competitive advantage [46]. Consequently, a clearly defined SPI program driven by strategic business needs, have been suggested as a key factor for success in SPI (e.g., [13], [17], [48], [55], [78], [98]). Therefore, we would expect that a successful SPI program is one, in which SPI goals and policies have been clearly aligned with business goals, identifying the nature and direction of any SPI actions.

Furthermore, learning theorists [18], [63], [96] have rejected context-free transfer models that isolate knowledge from practice and developed a view of learning as social construction, putting knowledge back into the contexts in which it has meaning. In other words, we expect that successful SPI align improvement activities to the real needs of the individual business rather than the abstract expectations of context-free models of SPI.

Thus, *we define business orientation as the extent to which SPI goals and actions are aligned with explicit and implicit business goals and strategies*. Hence, we argue that business orientation is a key factor to facilitate a successful SPI program. *Ceteris paribus*,

Hypothesis 1. *SPI success is positively associated with business orientation.*

2.2 Involved Leadership

Major change requires leadership. Thus, a predominant theme in the quality management [3], [15], [77], [82], [97] and SPI literature [44], [47], [55], [86], [98] is the importance of leadership commitment and involvement in the implementation of improvement actions. Such involvement is of paramount importance to SPI since top executives often are committed to the status quo [50], and also because they ultimately determine the priority of the organization's SPI program.

Creating a vision is considered a key element in most leadership frameworks [56], and those leading the organization must take an active role in describing a desired future and energize commitment to moving toward it. However, traditional views on leaders as "heroes" who set the direction, make key decisions, and energize the troops are changing [75], [84]. Creating vision is not enough; for assimilated learning to occur, leadership at any organizational level must engage in hands-on implementation of the vision. Thus, for building software organizations with learning capabilities, we need involved leadership at all levels that are committed to learning and SPI, and that take responsibility for enrolling others.

Accordingly, *we define involved leadership as the extent to which leaders at all levels in the organization are genuinely committed to and actively participate in SPI*. Hence, we argue that involved leadership is a key factor to facilitate a successful SPI program. *Ceteris paribus*,

Hypothesis 2. *SPI success is positively associated with involved leadership.*

2.3 Employee Participation

Employee participation and the way people are treated has been noted as a crucial factor in organizational management and development ever since Mayo's [68], [69] famous productivity studies at Western Electric's Hawthorne plant. The results of these studies started a revolution in management thinking, showing that even routine jobs can be improved if the workers are treated with respect. Besides, the Tavistock studies [92] reframed the view of organizational systems, proposing the *sociotechnical system* (STS) in which the best match would be sought between the requirements of the interdependent social *and* technical system. Together, these studies showed that there were powerful alternatives to the pervasive views of Taylor's concept of scientific management [90] and Weber's description of bureaucracy [95]. STS also has a particularly strong position in Scandinavia and the UK, which have a long tradition for work-place democracy and participation. In fact, many of the early experiments of STS were pioneered in Norway [91].

Since then, participation and involvement has been one of the most important foundations of organization development and change. Participation is also one of the fundamental ideas of TQM, and has always been a central goal and one of the pillars of organizational learning.

With SPI becoming increasingly important, and people increasingly recognized as the principal source of competitive advantage, software organizations need to encourage and support collaboration in work groups and project teams, and engage in organizational learning [22], [31], [43], [47], [86]. Participation is also fundamental for creativity and innovation and for showing respect for the opinions of others and for the belief that learning and competitive advantage can result from an organizational climate that cultivates a diversity of ideas and opinions [67], [83].

Therefore, participation should be offered and managed in such a way as to allow all employees to improve their work and to feel a sense of contribution to the organization and its mission. In our view, then, SPI is neither top-down nor bottom-up—rather, it is participative at all levels. Hence, *we define employee participation as the extent to which employees use their knowledge and experience to decide, act, and take responsibility for SPI*. Consequently, we argue that employee participation is a key factor to facilitate a successful SPI program. *Ceteris paribus*,

Hypothesis 3. *SPI success is positively associated with employee participation.*

2.4 Concern for Measurement

Software measurement is widely recognized as an essential part of understanding, controlling, monitoring, predicting, and evaluating software development and maintenance projects [32], [45], [49], [93] and as a necessary part of any SPI or change program [7], [48], [55], [61].

This position is by no means unique to the software community. Measurement and, in particular, measurement of customer satisfaction, is at the heart of quality management and is also a major concern for related disciplines such as organization development [26], organizational learning [5], strategic planning [60], and business process reengineering [28], [52].

Measurement and analysis is not only an efficient means for identifying, recommending, and evaluating process change; it can also be of crucial importance for assisting and guiding ongoing change [73]. A major concern for a measurement program, therefore, is to provide opportunities for developers to participate in analyzing, interpreting, and learning from the results of measurements and to identify concrete areas for improvement. There is, thus, reason to believe that providing a constant flow of quality data from a variety of sources may lead to improved performance.

Hence, *we define concern for measurement as the extent to which the software organization collects and utilizes quality data to guide and assess the effects of SPI activities*, and argue that concern for measurement is a key factor to facilitate a successful SPI program. *Ceteris paribus*,

Hypothesis 4. *SPI success is positively associated with concern for measurement.*

2.5 Learning Strategy

A critical challenge facing software organizations is the dilemma of maintaining the capabilities of both efficiency and flexibility. This situation, which requires the management of both stability and change, has led researchers and practitioners to distinguish between the more modest, or evolutionary, efforts toward change and those that are more fundamental and, in a sense, revolutionary (e.g., [5], [64]).

In other words, software organizations can engage in two broad kinds of learning strategies. They can engage in *exploitation*—the adoption and use of existing knowledge and experience, and they can engage in *exploration*—the search for new knowledge, either through imitation or innovation [36]. Exploitation involves improving existing capabilities by refining, standardizing, routinizing, and elaborating established ideas, paradigms, technologies, strategies, and knowledge.

In contrast, exploration involves learning through discovery and experimenting with ideas, paradigms, technologies, strategies, and knowledge in hope of finding new alternatives and untapped opportunities that are superior to current practice.

Finding a good balance between exploration and exploitation is a recurrent problem of theories of adaptation [65], which talk about balancing search and action, variation and selection, change and stability, and diversity and unity. A basic problem is thus to engage in enough exploitation to ensure short-term results and, concurrently, to engage in exploration to ensure long-term survival.

Based on the preceding discussion, we contend that exploitation and exploration are linked in an enduring symbiosis, and that each form requires the other in order to contribute effectively to a software organization's survival and prosperity. Consequently, *we define learning strategy as the extent to which a software organization is engaged in the exploitation of existing knowledge and in the exploration of new knowledge*. Hence, we argue that balancing the refinement of the existing skill base with the experimentation of new ideas is important in order to find alternatives that improve on old ones. *Ceteris paribus*,

Hypothesis 5. *SPI success is positively associated with exploitation of existing knowledge.*

Hypothesis 6. *SPI success is positively associated with exploration of new knowledge.*

2.6 Joint Contribution of Facilitating Factors

So far, we have discussed the possible independent contributions of business orientation, involved leadership, employee participation, concern for measurement, exploitation of existing knowledge, and exploration of new knowledge to SPI success. However, dependent variables such as organizational performance or other success measures, like the one used in this study, are rarely determined by one single independent variable. In addition to examining the independent contributions, therefore, we also want to examine the joint contribution of the independent variables to SPI success. *Ceteris paribus*,

Hypothesis 7. *The six independent variables of business orientation, involved leadership, employee participation, concern for measurement, exploitation of existing knowledge, and exploration of new knowledge will explain a large amount of the variance in SPI success.*

3 Research Method

3.1 Population and Sample

Software intensive organizations are considered as the target population for this study. This population includes companies of different sizes, developing either software or combined software and hardware products for a wide variety of markets.

A random sample of 154 software and quality managers in the Norwegian IT industry with corporate membership in the Association of the Norwegian Software Industry or the Norwegian IT Technology Forum were contacted by telephone to request participation in the study prior to mailing the questionnaires. Since the unit of analysis in this study was the software organization defined as a whole organization or an independent business unit within a larger company, the managers were asked to answer on behalf of their respective organizations.

We provided the respondents with self-addressed, stamped return envelopes. Also, by keeping the questionnaire as short as possible (a pilot study of the survey questionnaire showed that respondents needed about 10 minutes to complete it), we combined several well-proven techniques for improving the response rate of mailed questionnaires.

A total of 120 software and quality managers representing whole organizations or independent business units within 55 companies completed and returned the questionnaire. This is within the limits for both adequate statistical power and generalizability of the results in this study. Furthermore, this represents an effective response rate of 77.9 percent, which is well above the minimum norm of 40 percent for representatives of organizations and mid-level managers suggested for academic studies by Baruch [6]. Given the high response rate in this study, no further analysis was done on the differences between respondents and nonrespondents. Table 1 shows the characteristics of the survey sample.

3.2 Variables and Measures

Based on the definitions of the factors identified in the conceptual model (see Fig. 1), one dependent, six independent, and two moderating variables were operationalized and used to collect the data in this study.

Independent variables. To measure the extent to which each of the six independent variables were practiced, we used multi-item, five-point, bipolar Likert scales that ranged from "strongly disagree" (1) to "strongly agree" (5) for all indicators. The item ratings were summarized to form a summated rating scale for each independent variable. Furthermore, since this is the first study of its kind within SPI, all items were written specifically for this study.

Dependent variable. We operationalized and measured SPI success based on two multi-item measures. Each manager was asked to rate, on 5-point bipolar Likert scales, 1) the level of perceived SPI success and 2) the performance of their organization for the past three years with respect to cost reduction, cycle time reduction, and customer satisfaction. Two items were used to measure the level of perceived SPI success, while three items were used to measure organizational performance. As for the independent variables, all items were written specifically for this study. The ratings for the two performance dimensions were averaged to form a single measure of overall SPI success.

Moderating variables. Two moderating variables—*environmental conditions* and *organizational size*—were operationalized and included in the study to capture the most influential sources of variation in software organizations [36], [38]. Environmental conditions were measured using two semantic differential items, which were rated on a seven-point, bipolar, adjectival, graphic scale. The two items were stable versus unstable environment and predictable versus unpredictable environment. Organizational size was defined as the number of software developers in the organization.

The Appendix shows the summated rating scale items for the independent and dependent variables and the moderating variable environmental conditions along with the corresponding reliability measures (Cronbach's α).

3.3 Data Collection Procedure

The questionnaire consisted of two parts: The first part asked for managers' ratings of their software organizations with respect to factors of SPI success and organizational performance, and the second part asked for general background information and an assessment of the environment.

In the first part, 36 separate items were used to measure the six factors of SPI success, while five items were used to measure the performance construct. Thus, each manager generated seven scores; one for each of the critical factors, and one for the success measure.

In the second part of the questionnaire, each manager assessed their organization's environment in addition to providing general background and demographic information.

TABLE 1
Characteristics of the Survey Sample

Characteristics of the Respondents	**Mean**	**S.D.**
Average #years in the company	8.4	6.5
Average #years with software development	11.4	7.4
Highest completed education	**Freq.**	**Percent**
Bachelor's degree	38	31.7%
Master's degree	74	61.7%
Doctoral degree	5	4.2%
Other	3	2.5%
Job function		
Software manager	95	79.2%
Quality manager	25	20.8%
Characteristics of the Respondents' Company		
Number of software developers		
Less than or equal to 30	45	37.5%
Between 30 and 200	31	25.8%
More than or equal to 200	44	36.7%
Primary industry group		
Public sector	7	5.8%
Banking/finance/insurance	12	10.0%
Manufacturing	21	17.5%
IT sector	68	56.7%
Other	12	10.0%
Type of product business		
Standard applications (shelfware)	31	25.8%
Tailor made solutions, external customers	84	70.0%
Tailor made solutions, internal customers	5	4.2%
Quality system in use		
Yes	86	71.7%
No	34	28.3%

3.4 Assessment of Reliability and Validity

Reliability refers to the consistency and stability of a score from a measurement scale. The reliability of the multiple-item measurement scales was evaluated by internal consistency analysis, using coefficient alpha [23] and detailed item analysis based on Nunnally's method [76]. We analyzed the correlation of items within each scale (item-item), the corrected item-to-total (item-scale) correlations, the effects of reliability if the item was deleted, and the item standard deviation scores to determine which items were candidates for deletion from the scale.

This analysis revealed that to obtain satisfactory values for coefficient alpha while retaining the domain coverage only required one item to be eliminated from the exploitation scale (marked with an asterisk in the Appendix). Table 2 reports the original sets of measurement items associated with the key factors, the items dropped from the original sets to increase alpha, and the reliability coefficients for the resulting scales. Furthermore, Table 2 shows that the

TABLE 2
Reliability Analysis

Independent variables	**Item numbers**	**Number of items**	**Items deleted**	α
1. Business Orientation	1-5	5	none	.81
2. Involved leadership	6-10	5	none	.87
3. Employee participation	11-17	7	none	.80
4. Measurement	18-23	6	none	.81
5. Exploitation	24-29	6	no. 29	.78
6. Exploration	30-37	8	none	.85

TABLE 3
Summary of Factor Matrices for Each Construct

Independent variables	Eigen-value	Item loading range	#items with loadings > 0.6
1. Business Orientation	2.9	.72 to .82	5 (out of 5)
2. Involved leadership	3.3	.78 to .89	5 (out of 5)
3. Employee participation	3.2	.62 to .72	7 (out of 7)
4. Measurement	3.1	.43 to .83	4 (out of 6)
5. Exploitation	2.7	.65 to .84	5 (out of 5)
6. Exploration	3.9	.61 to .78	8 (out of 8)

reliability coefficients ranged from 0.78 to 0.87. Since reliability coefficients of 0.7 or higher are considered satisfactory [76], all scales developed for this study were judged to be reliable. This is good for an instrument that is composed entirely of new scales, particularly since the scales do not contain large numbers of items.

Three kinds of validity are of special concern for this study; content validity, construct validity, and criterion-related validity. *Content validity* has to do with the degree to which the scale items represent the domain of the concept under study. Our procedure for ensuring content validity followed the general recommendations of Cronbach [24] and Straub [89], and included:

1. an exhaustive search of the literature for all possible items to be included in the scales,
2. an exploratory study in representative companies to find possible items and scales,
3. review of the proposed scales by experts of both psychological testing and SPI, and
4. pilot test of the scales on a set of respondents similar to the target population.

Hence, we argue that our six measurement scales representing the facilitating factors of SPI success developed in this study have content validity since selection of measurement items was based on generally accepted procedures to obtain content validation [4].

Construct validity is an operational concept that examines whether the measurement scales represent and act like the attributes being measured. Assuming that the total score of a scale is valid, the extent to which an individual item measures the same thing as the total score is an indicator of the validity of that item. Hence, construct validity of the measurement scales was evaluated by confirmatory factor analysis.

The construct validity of the six measurement scales was evaluated by analyzing the items of each scale using principal components analysis with VARIMAX rotation. We used a combination of eigenvalues [59], cut-off points of the scree plots [19], and factor loadings [21] as a guide for interpreting the dimensionality of the scales. Table 3 shows the eigenvalues and item loading ranges for each scale. Analysis of the eigenvalues showed that five of the six scales formed a single factor. In the case of the exploration scale, two components seemed to emerge with eigenvalues greater than 1.0. However, the eigenvalue of the second factor was only slightly above this threshold (1.02). Furthermore, the scree plot showed a clear break after the first component. In addition, all item loadings for this component were greater than 0.6. This is in accordance with Stevens [88] who suggested that a reliable factor must have four or more loadings of at least 0.6 when the number of cases is below 150. These results indicate that all scales achieved a high degree of unidimensionality and, hence, construct validity.

Criterion-related validity is concerned with the degree to which the scales under study are related to an independent measure of the relevant criterion.

The criterion-related validity of the measurement instrument was evaluated by computing the multiple correlation (R) between the measures of the six independent variables and SPI success. The multiple correlation coefficient was 0.76. Cohen [20] suggested that a multiple correlation coefficient of 0.14 corresponds to a small effect size, that coefficients of 0.36 correspond to a medium effect size, and that coefficients above 0.51 correspond to a large effect size. Thus, we conclude that the independent variables have a high degree of criterion-related validity.

All in all, then, the results of reliability and validity analyses showed that the instrument has desirable psychometric properties. A more detailed analysis and discussion of the validity and reliability of these variables is presented in [35].

3.5 Data Analysis Techniques

We used parametric statistics such as the Pearson product-moment correlation coefficient, the t-test, the F statistic, and multiple regression analysis to analyze the data in this study. Using such parametric tests for summated rating scales, which are not strictly interval, does not lead, except in extreme cases, to wrong statistical decisions [16], [76], [94]. An important prerequisite, therefore, is that the assumptions of these analyses be met.

The results of testing these assumptions showed that kurtosis, skewness, and the one-sample Kolmogorov-Smirnov tests for all variables were within the acceptable range for the normal distribution assumption. Also, the assumptions of homoscedasticity, linearity, and independence of the error terms were supported, and no influential

TABLE 4
Item Means, Standard Deviations, and Correlations among the Independent Variables

Independent variable (*N* = 120)	Mean	S.D.	1	2	3	4	5	6
1. Business orientation	3.27	0.66	1.00					
2. Involved leadership	3.52	0.69	0.63***	1.00				
3. Employee participation	3.49	0.52	0.39***	0.37***	1.00			
4. Measurement	3.26	0.61	0.45***	0.41***	0.24**	1.00		
5. Exploitation	3.34	0.62	0.57***	0.46***	0.47***	0.41***	1.00	
6. Exploration	3.43	0.54	0.17*	0.11	0.46***	0.13	0.18*	1.00

Notes: * $p < 0.05$ ** $p < 0.005$ *** $p < 0.0005$

1. All *t*-tests are one-tailed.

observations were identified. Furthermore, investigations of collinearity and multicollinearity indicated no problems.

In other words, *there were no extreme violations to the basic assumptions underlying the chosen data analysis techniques* that could justify the use of less powerful nonparametric statistics. All quantitative analyses were conducted using SPSS.

4 Results

Table 4 shows the means, standard deviations, and correlations among the independent variables. Out of 15 correlations between the independent variables, two have a correlation coefficient larger than 0.5. The highest correlation (0.63) is between involved leadership and business orientation.

4.1 Testing Individual Relationships

Hypotheses 1 through 6 consider the individual relationships between SPI success and each of the six independent variables. The testing of these hypotheses calls for the use of bivariate correlations. In addition, to examine each independent variable's correlation with overall SPI success, we also examined the correlations with each of the two underlying success measures: perceived level of success and organizational performance. However, appropriate tests of the bivariate correlations require that the two contextual factors, environment and organizational size, be partialled from the analysis. Table 5 shows both the zero-order correlations and the partial correlations when the effects of the contextual variables have been removed between the independent variables and each of the SPI success measures.

The zero-order Pearson correlation results between exploration of new knowledge and overall SPI success showed a positive and significant correlation coefficient of 0.21 ($p < 0.05$). The zero-order correlations with perceived level of success ($r = 0.15, p < 0.05$) and organizational performance ($r = 0.20, p < 0.05$) were also positive and significant. Furthermore, all partial correlations were positive and significant, ranging from $pr = 0.18 (p < 0.05)$ to $pr = 0.25 (p < 0.005)$.

TABLE 5
Tests of Hypotheses H1-H6

	Overall SPI success		Perceived level of success		Organizational performance	
Independent variable (*N* = 120)	*r*	*pr*	*R*	*pr*	*r*	*pr*
1. Business Orientation	.62***	.61***	.59***	.59***	.46***	.44***
2. Involved leadership	.55***	.54***	.58***	.57***	.34***	.34***
3. Employee participation	.55***	.59***	.49***	.52***	.43***	.49***
4. Measurement	.50***	.48***	.44***	.44***	.40***	.38***
5. Exploitation	.59***	.59***	.55***	.56***	.44***	.44***
6. Exploration	.21*	.25**	.15*	.18*	.20*	.25**

Notes: * $p < 0.05$ ** $p < 0.005$ *** $p < 0.0005$

1. All *t*-tests are one-tailed.
2. Partial *r* (*pr*) is the correlation between the success measure and the independent variable when the contextual factors (environment and organizational size) are held constant.

TABLE 6
Regression Analysis of Overall SPI Success

Independent variables	*B*	Std. error	*Beta*	*t*-value
1. Business Orientation	.157	.056	.245	2.778**
2. Involved leadership	.076	.050	.125	1.526
3. Employee participation	.173	.045	.297	3.803***
4. Measurement	.115	.041	.200	2.828**
5. Exploitation	.121	.055	.177	2.201*
6. Exploration	-.019	.034	-.039	-.559

Notes: $R = .76$ $R^2 = .58$ Adj. $R^2 = .56$ ($F = 25.95$***)
*$p < .05$ **$p < .01$ ***$p < .0005$

Similarly, all zero-order correlation results between the remaining five independent variables and overall SPI success showed *large positive and highly significant* correlations, ranging from $r = 0.50 (p < 0.0005)$ to $r = 0.62 (p < 0.0005)$. In addition, all zero-order correlations with perceived level of success and organizational performance were positive and highly significant, ranging from $r = 0.34 (p < 0.0005)$ to $r = 0.59 (p < 0.0005)$. Furthermore, all partial correlations with overall SPI success were positive and highly significant, ranging from $pr = 0.48 (p < 0.0005)$ to $pr = 0.61 (p < 0.0005)$. Finally, all partial correlations with perceived level of success and organizational performance were positive and highly significant, ranging from $pr = 0.34 (p < 0.0005)$ to $pr = 0.59$ $(p < 0.0005)$.

Taken together, all zero-order and partial correlations involved in testing Hypotheses 1 through 6 were significant and in the hypothesized directions. This indicates support for the validity of all major measures used in this research. Thus, the findings in Table 5 support Hypotheses 1 through 6, along with the underlying assumption that SPI provides increased levels of performance to the organization.

4.2 Testing Overall Relationships

We examined the joint contribution of the independent variables to the explanation of SPI success as well as their contribution to the variate and its predictions in order to identify the critical factors for SPI success.

A "large" amount of the variance in SPI success was defined as a large effect size, $f^2 \geq 0.35$, according to Cohen's (1988) categorization, where:

$$f^2 = \frac{R^2}{1 - R^2}. \quad (1)$$

Given this relationship between effect size (f^2) and the squared multiple correlation (R^2), a large effect size of $f^2 \geq 0.35$ corresponds to a squared multiple correlation of $R^2 \geq 0.26$ and a multiple correlation coefficient of $R \geq 0.51$.

In order to make inferences about the multiple correlation, we used Konishi's [62] extension of the Fisher *r*-to-*Z* transformation to test the exact hypothesis that the population value of R is large. That is, we tested the modified null hypothesis

$$H7_0 : \rho < 0.51, \text{ against the alternative}$$
$$H7_A : \rho \geq 0.51.$$

The test statistic, C_R, is referred to the normal distribution, and given by:

$$C_R = \left\{ Z_R - \zeta_\rho - \frac{1}{2\rho(N-1)} \left(K - 1 + \rho^2\right) \right\} \sqrt{N-1} \quad (2)$$

in which, Z_R is the Fisher Z value corresponding to the sample R value, ζ_ρ is the Fisher Z value corresponding to the population ρ value stated in the null hypothesis, K is the number of predictor variables, and N is the sample size.

The test statistic for the modified null hypothesis with $K = 6$ predictor variables, a sample size of $N = 120$, and a sample multiple correlation of $R = 0.76$ was $C_R = 4.26$. In a standard normal distribution, a value of 4.26 is highly significant and, thus, falls outside the acceptance region for the $\alpha = 0.05$ level. Hence, in light of the sample evidence, we rejected the null hypothesis that the true value of the multiple correlation is less than 0.51, and accepted the alternate hypothesis that *the six predictor variables significantly* $(p < 0.00005)$ *explain a large amount* $(\rho \geq 0.51)$ *of the variance in SPI success*.

The contributions of each independent variable to the variate, however, differ substantially. As can be seen in Table 6, two of the independent variables—*involved leadership and exploration of new knowledge*—did not contribute significantly to the explanation of SPI success. Of the remaining four significant variables, *employee participation* seems to be associated with the highest explanatory power, since it achieved the highest standardized regression coefficient $(\beta = 0.297, t = 3.80, p < 0.0005)$. Next came *business orientation* $(\beta = 0.245, t = 2.78, p < 0.01)$, followed by *concern for measurement* $(\beta = 0.200, t = 2.83, p < 0.01)$ and, finally, *exploitation of existing knowledge* $(\beta = 0.177, t = 2.20, p < 0.05)$.

It is important to note, however, that the weights the independent variables carry in the multiple regression equation are always relative to the entire set of predictors employed. That is, the predictive strength that some independent variable seems to show for the dependent variable may differ if the set of remaining variables is altered.

The two contextual factors were added to the model to test the sensitivity of the variate to organizational size and environmental turbulence. The estimated coefficients of the independent variables of the original model were not significantly changed by the introduction of the contextual factors. *The results hold for small, as well as large, software organizations, and for organizations operating in stable as well as in turbulent environments*.

These results support the robustness of the original regression model results. Furthermore, the model with the two contextual factors added ($R\text{-square} = 0.59$, adjusted $R\text{-square} = 0.56, F = 19.60, p < 0.0005$) did not explain significantly more variance as compared with the original model ($R\text{-square}=0.58$, adjusted $R\text{-square}=0.56, F=25.95, p < 0.0005$).

In conclusion, the multiple regression analysis did not support Hypotheses 2 and 6, provided that all hypotheses are considered as one set. In other words, considering both bivariate correlation analyses and the multiple regression analysis, we find strong support for hypotheses 1, 3, 4, 5, and 7, and partial support for hypotheses 2 and 6.

5 Discussion

A quantitative survey was performed to investigate the key factors of success in SPI. The results of the bivariate correlational and multiple regression analyses showed that SPI success depends critically on six organizational factors. In the following, we discuss the insights that can be gained for practitioners and researcher from these findings, the limitations of the study, and some suggestions for further research.

5.1 Practitioner and Researcher Insights

Organizational knowledge is created by ongoing experience and experimentation of organizational members. Without individuals who learn, there can be no improvement. The fundamental role of managers and leaders is, thus, to shape and create organizational contexts that are supportive of and conducive to organizational learning and SPI.

First of all, *business orientation*, that is, the extent to which SPI goals and actions are aligned with explicit and implicit business goals and strategies, was identified as one of the factors with the strongest influence on SPI success. This supports the hypothesis that SPI success is positively associated with business orientation.

This finding is important, and suggests that both *practitioners and researchers should direct significant effort toward understanding shared domain knowledge between software and business executives*. However, there are two basic conditions for establishing such connections and making communications between these groups effective. First, each group must respect the expertise of the other, and must acknowledge the relevance of that expertise to their own problems. Second, each group must have sufficient knowledge and understanding of the other groups' problems to be able to communicate effectively about them.

It is important to remember that experience has shown that such shared understandings are unlikely to occur unless a sufficient number of members of each group have had actual experience with the activities and responsibilities of the other group [85]. This suggests that the software organization should be active in creating the possibilities for such connections to be formed, for example, through personnel rotation programs, and promotion and recruitment strategies.

The critical importance of a business orientation in SPI suggests that software organizations should focus on both short-term and long-term alignment of SPI goals with business goals. They should focus attention on how to achieve a high degree of mutual understanding of current objectives as well as on how to achieve congruence of long-term SPI goals with business strategies.

Thus, the importance of aligning SPI goals with business goals implies that if the SPI strategy is not kept in line with changes in the business strategy, there is a risk that the organization's software processes may end up as a burden rather than advance the business. Therefore, our results also suggest that understanding the business and organizational context is critical for achieving alignment between SPI activities and business strategy, which in turn is of paramount importance for the success of the SPI program. Similarly, by viewing the organization as a living organism, Zahran compared the introduction of a new process with transplanting a new organ [98]; if the new process is not aligned with the business strategy or does not match the organization's culture, it will be rejected by the "organizational" body.

Involved leadership—the extent to which leaders at all levels in the organization are genuinely committed to and actively participate in SPI—had a strong and highly significant correlation with overall SPI success, which supports the hypothesis that SPI success is positively associated with involved leadership.

A surprising result, however, was the insignificant importance of involved leadership in predicting SPI success. All the main studies on which this investigation is based share a strong belief in the importance of management commitment for the successful implementation of SPI. Furthermore, the quality management literature and the organizational learning literature also seem to share a strong belief in the importance of management commitment for improving organizational performance. With this background, it is surprising, both from a theoretical and practical perspective, that involved leadership, defined as the extent to which leaders at all levels in the organization are genuinely committed to and actively participate in SPI, is a non-significant predictor of SPI success. On the other hand, the results are in agreement with Abrahamsson's findings that many SPI initiatives do not require management commitment beyond obtaining the resources needed [1]. There could be several explanations for this seemingly surprising result.

First, there is no common understanding of the role of involved leadership or management commitment in SPI in terms of conceptual definitions and operationalized measures [2], [35]. Therefore, we cannot know what is actually meant by, e.g., management commitment in a particular study and it will be difficult (if not impossible) to compare the results of one study with those of another. So, until SPI researchers can agree upon the conceptual definitions used to describe the phenomenon under study, we cannot go beyond a face validity perspective and truly tell whether findings are comparable or not.

Second, we found that multicollinearity should not have a substantial impact on the estimated regression variate in this investigation. However, the results of the multiple regression analysis show that multicollinearity, nevertheless, does have an impact on the composition of the variate, and that correlations among the independent variables may make some variables redundant in the predictive effort. This is the case for involved leadership and also for exploration of new knowledge. However, this does not reflect their individual relationships with the dependent variable. Instead, it indicates that in a multivariate context, they are not needed together with the remaining set of four independent variables to explain the variance in SPI success. Therefore, we cannot determine the importance of involved leadership based solely on the derived variate, since relationships with the other independent variables may "mask" relationships that are not needed for predictive purposes, but represent substantive explanatory findings nonetheless.

Third, the role of management is often to ensure that SPI goals and actions are closely aligned with business goals and strategies, which is indicated by the highly significant correlation coefficient of 0.63 in Table 4 between business orientation and involved leadership. This suggests that involved leadership is important through business orientation. This is also in agreement with Grady's findings from his longitudinal studies in Hewlett-Packard that business aspects in the organization's SPI efforts had a strong influence on management commitment [48]. Furthermore, Debou and Kuntzmann-Combelles argued that management commitment implies business orientation and vice versa [29].

Finally, a large part of the published, and thus reviewed, studies on quality management, organizational learning, and SPI are from the US. There are, however, important cultural differences between the US and Europe in general [72] and between the US and Scandinavia in particular [41]. The Norwegian Work Environment Act, for example, stipulates that the employer has to negotiate with the employees before making "major changes" in production. Also, in terms of Hofstede's model (see [14]), the relatively high degree of workplace democracy and sociotechnical tradition in Scandinavian countries indicate that they might have a smaller power distance and a more collectivist culture than the US. This suggests a higher importance of employee participation and a correspondingly minor importance of involved leadership in predicting SPI success in these countries. This is also indicated by the results of the multiple regression analysis that showed that together with business orientation, employee participation was the most important factor in predicting SPI success.

Thus, together with business orientation, *employee participation*, i.e., the extent to which employees use their knowledge and experience to decide, act, and take responsibility for SPI, was identified as the factor with the strongest influence on SPI success. This supports the hypothesis that SPI success is positively associated with employee participation. This is not surprising and it supports previous research on the role of employee participation in organizational management and development [26], [46].

Also, a multiple case study by Conradi and Dybå to explore the diffusion of knowledge and experience in software organizations [22], strongly support the importance of participation as a key factor for SPI success. Similarly, a multiple case study of tailor made assessments to focus software process improvement activities, showed that a structured process emphasizing participation in each and every step was a key factor for success [39].

This suggests that *people tend to support what they have participated in creating*, or to use Berger and Luckmann's words [11]: "It is more likely that one will deviate from programmes set up for one by others than from programmes that one has helped establish oneself." Thus, we learn a great deal more from our own experience than we do from those who are experienced.

Concern for measurement—the extent to which the software organization collects and utilizes quality data to guide and assess the effects of SPI activities—had a strong and highly significant correlation with overall SPI success. This supports the hypothesis that SPI success is positively associated with concern for measurement and suggests that *significant progress in software development depends on an ability to measure aspects of the process and the resulting artifacts, and to make analytical judgments based on the results of these measurements*. Thus, a lack of measurement leads to a lack of empirical validation of software techniques and ultimately to an inability of the organization to evaluate new (or old) ideas and to separate sound and useful practices from current fads.

Traditionally, measurement systems are used to provide upper management with data for decision making, and to assess the effects of organizational actions. Although this is still important, our concept of measurement not only includes these traditional uses; it also includes availability and feedback of data, and the use of data to guide SPI actions. So, while measurement in itself can have a significant effect on organizational behavior [73], our results suggest that the most effective use of data for organizational learning and SPI is to feed the data back in some form to the organization's members. Such feedback regarding actual performance not only motivates change in the behavior of individuals, groups, and organizations; it can also guide change in a specific direction. However, in order for feedback to change behavior through these mechanisms, the data must be perceived as valid and accurate. Also, the conditions surrounding the feedback process must support the nonthreatening use of data for identifying and solving problems.

This suggests that without clear and consistent feedback to help monitor software teams' progress toward their goals and the effects of their actions, it is difficult to learn. Besides, our results demonstrate that measurement is meaningless without interpretation and judgment by those who will make decisions and take actions based on them. Within the context of SPI, therefore, *it is important that measurement systems are* *designed by software developers for learning, rather than by management for control*. Also, an important observation is that a few measures that are directly related to the software process are better than a multitude of measures that produce a lack of focus and confusion about what is important and what is not [40], [80].

Our findings thus suggest that data-based methods in general and feedback, in particular, are effective tools for the successful application of SPI, and that knowledge about *how things are* can be a potential force that moves software organizations toward *how things should be*.

Two fundamentally different learning strategies were identified as part of our investigation: *exploitation* of existing knowledge and *exploration* of new knowledge. Consequently, we defined learning strategy as the extent to which a software organization is engaged in the exploitation of existing knowledge and in the exploration of new knowledge. Both learning strategies had a significant correlation with overall SPI success, which supports the hypotheses that SPI success is positively associated with the exploitation of existing knowledge as well as with the exploration of new knowledge.

However, as in the case of involved leadership, exploration of new knowledge had an insignificant importance in *predicting* SPI success. We see two plausible explanations for this. First, as we explained with respect to involved leadership, relationships with the other independent variables may "mask" relationships that are not needed for predictive purposes, but represent substantive explanatory findings nonetheless. Therefore, we cannot determine the importance of exploration of new knowledge based solely on the derived variate.

Second, our concept of exploration includes issues such as innovation and creativity, questioning of "established" truths, flexibility, minimum critical specification, and diversity, which, within a sociotechnical perspective, can be seen as closely related to conceptions of participation. This is also indicated by the highly significant correlation coefficient of 0.46 in Table 4 between employee participation and exploration of new knowledge, which suggests that exploration of new knowledge is important through employee participation.

With respect to the two contextual factors, we found that the general results hold for small, as well as large, software organizations, and for organizations operating in stable as well as in turbulent environments. Still, a closer examination of organizational size showed that large successful and small successful organizations differed fundamentally in their respective approach to SPI, specifically with respect to participation and their preferred mode of learning. Small successful organizations reported higher levels of employee participation and exploration of new knowledge than the larger organizations (see [38]). This suggests that both modes of learning are important for successful SPI. Moreover, it suggests that *software organizations should balance their learning efforts between exploitation and exploration*.

It is also interesting to note that there was no difference in the level of exploitation between small and large organizations regardless of the environment, while there was a marked difference in the level of exploration. The results showed that small software organizations engaged in significantly more exploration in turbulent environments than large software organization (see [36]), which suggests that *the main difference between small and large software organizations is the ways in which they react to unstable and changing stimulus situations*.

5.2 Limitations

Although we have discussed several implications of our results, the research reported here is not without its limitations. First, the results have been discussed as though the facilitating factors caused SPI success. However, the investigation was cross-sectional and these assumptions of causality are not technically justified. It is possible, for example, that there are complex feedback mechanisms by which performance in one time period is affected by performance in previous periods [66]. Both success and lack of success might very well be self-reinforcing. On the other hand, there could also be negative feedback mechanisms by which SPI success or failure creates countervailing tendencies. It is possible therefore that performance below target levels increases organizational efforts and, thus, the likelihood of subsequent success. Accordingly, success can also be seen as a trigger for adjustments in the opposite direction through decreased innovation or increased slack. Therefore, many of the facilitating factors that seem likely to influence SPI success can themselves be influenced by prior success or lack of success. Finding the "true" causal structure of SPI success based on incomplete information generated by prior experience is, therefore, problematic.

Second, although generally accepted psychometric principles were used to develop the measurement instrument, the variables were still measured on the basis of subjective performance definitions. Performance measures such as the return on investment, net present value, and payback periods are often regarded as objective measures. However, attempts to provide objective definitions of such measures may be as open to criticism as subjective definitions [12], [35], which points to a general problem of defining and measuring success in studies of SPI. The question, therefore, is not whether such measures are subjective or not, but what purpose they serve.

Finally, a further complication is that the independent variables were assessed using retrospective recall. This involves a risk for the introduction of retrospective bias. It is possible, therefore, that performance information itself colors subjective memories and perceptions of possible causes of SPI success or failure. Software organizations are constantly worried about their performance, and "common wisdom" has many explanations for good and poor performance. As a result, retrospective reports of independent variables may be less influenced by memory than by a reconstruction that connects common wisdom with the awareness of performance results [66].

Despite these limitations, this study contributes to the growing literature on empirical software engineering research and provides empirical support for the importance of organizational issues in SPI.

5.3 Future Work

SPI success measurement is a controversial issue and more research is needed to study it. Several levels of analysis are possible—e.g., individual, group, process, and organization—each with complex interactions with the others. Also, several, and possibly conflicting, dimensions (e.g. faster, better, and cheaper) and viewpoints (e.g., economic, engineering, organizational, and sociotechnical) are relevant. Further research should be related to the study of new and improved measures of SPI success, comparison of measurement instruments, and validation of SPI success measures.

In addition, context variables, such as organizational size and environmental turbulence, did not play an important role in *predicting* SPI success in this study. However, under other circumstances, the independent variables of the present study could act as moderating or mediating variables in other studies. Further research is therefore needed to investigate the importance of such variables to several types of SPI problems and to validate the approaches proposed for solving them. Such studies should combine factor research with process research in order to provide satisfactory levels of external validity. This way, we can tailor software engineering methods and improvement strategies to better help software organizations succeed with their unique business of software.

There are certainly other directions for further research. However, the value of any such future work depends on the specific goals of each particular investigation.

6 Conclusion

The study focused on identifying the key factors for success in SPI by a quantitative survey. The results indicate support for all of the hypotheses in the proposed model and demonstrate the existence of important factors for SPI success. From a theoretical perspective, these findings add an important new dimension to empirical software engineering research in that they verify the importance of organizational factors for SPI success. From a practical perspective, this suggests that, rather than trying to imitate technical procedures, software organizations should focus their SPI efforts on creating an organizational culture within which these procedures can thrive. This differs substantially from that found in most of the existing SPI literature, which focuses almost entirely on software engineering tools and techniques. Overall, from both theoretical and practical perspectives, an important new insight from this research is that organizational issues are at least as important in SPI as technology, if not more so.

Appendix

Measurement Instrument

A more detailed analysis and discussion of the validity and reliability of the operationalized measures for the variables in this study can be found in [35]. Cronbach's alpha is shown in the parenthesis for each measure.

Business Orientation ($\alpha = 0.81$)

1. We have established unambiguous goals for the organization's SPI activities.
2. There is a broad understanding of SPI goals and policy within our organization.
3. Our SPI activities are closely integrated with software development activities.
4. Our SPI goals are closely aligned with the organization's business goals.
5. We have a fine balance between short-term and long-term SPI goals.

Leadership Involvement ($\alpha = 0.87$)

6. Management is actively supporting SPI activities.
7. Management accepts responsibility for SPI.
8. Management considers SPI as a way to increase competitive advantage.
9. Management is actively participating in SPI activities.
10. SPI issues are often discussed in top management meetings.

Employee Participation ($\alpha = 0.80$)

11. Software developers are involved to a great extent in decisions about the implementation of their own work.
12. Software developers are actively contributing with SPI proposals.
13. Software developers are actively involved in creating routines and procedures for software development.
14. We have an ongoing dialogue and discussion about software development.
15. Software developers have responsibility related to the organization's SPI activities.
16. Software developers are actively involved in setting goals for our SPI activities.
17. We have an ongoing dialogue and discussion about SPI.

Concern for Measurement ($\alpha = 0.81$)

18. We consider it as important to measure organizational performance.
19. We regularly collect quality data (e.g. defects, timeliness) from our projects.
20. Information on quality data is readily available to software developers.
21. Information on quality data is readily available to management.
22. We use quality data as a basis for SPI.
23. Our software projects get regular feedback on their performance.

Exploitation of Existing Knowledge ($\alpha = 0.78$)

24. We exploit the existing organizational knowledge to the utmost extent.
25. We are systematically learning from the experience of prior projects.
26. Our routines for software development are based on experience from prior projects.
27. We collect and classify experience from prior projects.

28. We put great emphasis on internal transfer of positive and negative experience.
29. To the extent we can avoid it, we do not take risks by experimenting with new ways of working.*

Exploration of New Knowledge ($\alpha = 0.85$)

30. We are very capable at managing uncertainty in the organization's environment.
31. In our organization, we encourage innovation and creativity.
32. We often carry out trials with new software engineering methods and tools.
33. We often conduct experiments with new ways of working with software development.
34. We have the ability to question "established" truths.
35. We are very flexible in the way we carry out our work.
36. We do not specify work processes more than what are absolutely necessary.
37. We make the most of the diversity in the developer's skills and interests to manage the variety and complexity of the organization's environment.

Organizational Performance ($\alpha = 0.76$)

1. Our SPI work has substantially increased our software engineering competence.
2. Our SPI work has substantially improved our overall performance.
3. Over the past three years, we have greatly reduced the cost of software development.
4. Over the past three years, we have greatly reduced the cycle time of software development.
5. Over the past three years, we have greatly increased our customers' satisfaction.

Environmental Conditions ($\alpha = 0.80$)

1. Stable □ 1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 Unstable
2. Predictable □ 1 □ 2 □ 3 □ 4 □ 5 □ 6 □ 7 Unpredictable

*Starred items were removed from the final instrument and should not be used.

ACKNOWLEDGMENTS

This work was supported in part by the Research Council of Norway under Grant 118206/221. The author wishes to thank all respondents of the survey for their willingness to participate in the inquiries.

REFERENCES

[1] P. Abrahamsson, "Is Management Commitment a Necessity after All in Software Process Improvement?" *Proc. 26th Euromicro Conf.,* vol. 2, pp. 246-253, Sept. 2000.
[2] P. Abrahamsson, "Commitment Development in Software Process Improvement: Critical Misconceptions," *Proc. 23rd Int'l Conf. Software Eng. (ICSE 2001),* pp. 71-80, 2001.
[3] S.L Ahire, D.Y. Golhar, and M.A. Waller, "Development and Validation of TQM Implementation Constructs," *Decision Sciences,* vol. 27, no. 1, pp. 23-56, 1996.
[4] A. Anastasi and S. Urbina, *Psychological Testing,* seventh ed. Upper Saddle River, New Jersey: Prentice-Hall, 1997.
[5] C. Argyris and D.A. Schön, *Organizational Learning II: Theory, Method, and Practice,* Reading, Mass.: Addison-Wesley, 1996.
[6] Y. Baruch, "Response Rate in Academic Studies—A Comparative Analysis," *Human Relations,* vol. 52, no. 4, pp. 421-438, 1999.
[7] V.R. Basili and G. Caldiera, "Improve Software Quality by Reusing Knowledge and Experience," *Sloan Management Rev.,* vol. 37, no. 1, pp. 55-64, Autumn 1995.
[8] V.R. Basili and H.D. Rombach, "The TAME Project: Towards Improvement-Oriented Software Environments," *IEEE Trans. Software Eng.,* vol. 14, no. 6, pp. 758-773, 1988.
[9] V.R. Basili, F.E. McGarry, R. Pajerski, and M.V. Zelkowitz, "Lessons Learned from 25 Years of Process Improvement: The Rise and Fall of the NASA Software Engineering Laboratory," *Proc. 24th Int'l Conf. Software Eng. (ICSE 2002),* pp. 69-79, 2002.
[10] M. Beer, R.A. Eisenstat, and B. Spector, "Why Change Programs Don't Produce Change," *Harvard Business Rev.,* vol. 68, no. 6, pp. 158-166, 1990.
[11] P.L. Berger and T. Luckmann, *The Social Construction of Reality: A Treatise in the Sociology of Knowledge.* Harmondsworth: Penguin Books, 1966.
[12] M. Berry and R. Jeffery, "An Instrument for Assessing Software Measurement Programs," *Empirical Software Eng.,* vol. 5, no. 3, pp. 183-200, Nov. 2000.
[13] M. Biró and C. Tully, "The Software Process in the Context of Business Goals and Performance," *Better Software Practice for Business Benefit: Principles and Experience,* R. Messnarz and C. Tully, eds. CS Press, pp. 15-27, 1999.
[14] M. Biró, R. Messnarz, and A.G. Davison, "The Impact of National Cultural Factors on the Effectiveness of Process Improvement Methods: The Third Dimension," *ASQ Software Quality Professional,* vol. 4, no. 4, pp. 34-41, 2002.
[15] S.A. Black and L.J. Porter, "Identification of the Critical Factors of TQM," *Decision Sciences,* vol. 27, no. 1, pp. 1-21, 1996.
[16] L.C. Briand, K. El Emam, and S. Morasca, "On the Application of Measurement Theory in Software Engineering," *Empirical Software Eng.,* vol. 1, no. 1, pp. 61-88, 1996.
[17] J.G. Brodman and D.L. Johnson, "Return on Investment (ROI) from Software Process Improvement as Measured by US Industry," *Software Process Improvement and Practice,* pp. 35-47, 1995.
[18] J.S. Brown and P. Duguid, "Organizational Learning and Communities of Practice: Toward a Unified View of Working, Learning, and Innovation," *Organization Science,* vol. 2, no. 1, pp. 40-57, 1991.
[19] R.B. Cattell, "The Scree Test for the Number of Factors," *Multivariate Behavioral Research,* vol. 1, pp. 245-276, 1966.
[20] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences,* second ed. Hillsdale, New Jersey: Laurence Erlbaum, 1988.
[21] A.L. Comrey and H.B. Lee, *A First Course on Factor Analysis,* second ed. Hillsdale, N.J.: Erlbaum, 1992.
[22] R. Conradi and T. Dybå, "An Empirical Study on the Utility of Formal Routines to Transfer Knowledge and Experience," *Proc. Joint Eighth European Software Eng. Conf. (ESEC) and Ninth ACM SIGSOFT Int'l Symp. Foundations of Software Eng. (FSE),* Sept. 2001.
[23] L.J. Cronbach, "Coefficient Alpha and the Internal Consistency of Tests," *Psychometrica,* vol. 16, pp. 297-334, 1951.
[24] L.J. Cronbach, "Test Validation," *Educational Measurement,* R.L. Thorndike, ed., second ed., Washington: Am. Council on Education, pp. 443-507, 1971.
[25] P.B. Crosby, *Quality is Still Free: Making Quality Certain in Uncertain Times.* New York: McGraw-Hill, 1996.
[26] T.G. Cummings and C.G. Worley, *Organization Development and Change,* eighth ed. Cincinnati, Ohio: South-Western College Publishing, 2004.
[27] M.K. Daskalantonakis, "A Practical View of Software Measurement and Implementation Experiences within Motorola," *IEEE Trans. Software Eng.,* vol. 18, no. 11, pp. 998-1010, 1992.
[28] T.H. Davenport, *Process Innovation: Reengineering Work through Information Technology.* Boston, Mass.: Harvard Business School Press, 1993.
[29] C. Debou and A. Kuntzmann-Combelles, "Linking Software Process Improvement to Business Strategies: Experience from Industry," *Software Process: Improvement and Practice,* vol. 5, pp. 55-64, 2000.
[30] C. Debou, D. Courtel, H.-B. Lambert, N. Fuchs, and M. Haux, "Alcatel's Experience with Process Improvement," *Better Software Practice for Business Benefit: Principles and Experience,* R. Messnarz and C. Tully, eds., CS Press, pp. 281-301, 1999.

[31] C. Deephouse, T. Mukhopadhyay, D.R. Goldenson, and M.I. Kellner, "Software Processes and Project Performance," *J. Management Information Systems,* vol. 12, no. 3, pp. 187-205, 1996.

[32] T. DeMarco, *Controlling Software Projects: Management, Measurement and Estimation.* New York: Yourdon Press, 1982.

[33] W.E. Deming, *Out of the Crisis.* Cambridge, Mass.: MIT Center for Advanced Eng. Study, 1986.

[34] R. Dion, "Process Improvement and the Corporate Balance Sheet," *IEEE Software,* vol. 10, no. 4, pp. 28-35, 1993.

[35] T. Dybå, "An Instrument for Measuring the Key Factors of Success in Software Process Improvement," *Empirical Software Eng.,* vol. 5, no. 4, pp. 357-390, 2000.

[36] T. Dybå, "Improvisation in Small Software Organizations," *IEEE Software,* vol. 17, no. 5, pp. 82-87, Sept.-Oct. 2000.

[37] T. Dybå, "A Dynamic Model of Software Engineering Knowledge Creation," *Managing Software Eng. Knowledge,* A. Aurum et al., eds., Berlin: Springer Verlag, pp. 95-117, 2003.

[38] T. Dybå, "Factors of Software Process Improvement Success in Small and Large Organizations: An Empirical Study in the Scandinavian Context," *Proc. Joint Ninth European Software Eng. Conf. (ESEC) and 11th SIGSOFT Symp. Foundations Software Eng. (FSE-11),* pp. 148-157, 2003.

[39] T. Dybå and N.B. Moe, "Rethinking the Concept of Software Process Assessment," *Proc. European Software Process Improvement Conf. (EuroSPI '99),* Oct. 1999.

[40] T. Dybå, T. Dingsøyr, and N.B. Moe, *Process Improvement in Practice: A Handbook for IT Companies,* The Kluwer Int'l Series in Software Eng., Boston: Kluwer Academic Publishers, 2004.

[41] P. Ehn, "Scandinavian Design: On Participation and Skill," *Usability—Turning Technologies into Tools,* P.S. Adler and T.A. Winograd, eds., New York: Oxford Univ. Press, pp. 96-132, 1992.

[42] *SPICE: The Theory and Practice of Software Process Improvement and Capability Determination,* K. El Emam et al., eds., CS Press, 1998.

[43] K. El Emam, P. Fusaro, and B. Smith, "Success Factors and Barriers for Software Process Improvement," *Better Software Practice for Business Benefit: Principles and Experience,* R. Messnarz and C. Tully, eds., CS Press, pp. 355-371, 1999.

[44] K. El Emam, D.R. Goldenson, J. McCurley, and J. Herbsleb, "Modeling the Likelihood of Software Process Improvement: An Exploratory Study," *Empirical Software Eng.,* vol. 6, no. 3, pp. 207-229, 2001.

[45] N.E. Fenton and S.H. Pfleeger, *Software Metrics: A Rigorous and Practical Approach.* London: Int'l Thomson Computer Press, 1996.

[46] W.L. French and C.H. Bell Jr., *Organization Development: Behavioral Science Interventions for Organization Improvement,* sixth ed. Upper Saddle River, New Jersey: Prentice-Hall, 1999.

[47] D.R. Goldenson and J.D. Herbsleb, "After the Appraisal: A Systematic Survey of Process Improvement, its Benefits, and Factors that Influence Success," Technical Report, CMU/SEI-95-TR-009, Carnegie Mellon Univ., Software Eng. Inst., 1995.

[48] R.B. Grady, *Successful Software Process Improvement.* Upper Saddle River, New Jersey: Prentice-Hall, 1997.

[49] R.B. Grady and D. Caswell, *Software Metrics: Establishing a Company-Wide Program.* Englewood Cliffs, New Jersey: Prentice-Hall, 1987.

[50] D.C. Hambrick, M.A. Geletkanycz, and J.W. Fredrickson, "Top Executive Commitment to the Status Quo: Some Tests for Its Determinants," *Strategic Management J.,* vol. 14, no. 6, pp. 401-418, 1993.

[51] M. Hammer, *Beyond Reengineering: How the Process-Centered Organization is Changing Our Work and Our Lives.* London: HarperCollins, 1996.

[52] M. Hammer and J. Champy, *Reengineering the Corporation: A Manifesto for Business Revolution.* New York: Harper Business, 1993.

[53] W.L. Hays, *Statistics,* fifth ed. New York: Harcourt Brace, 1994.

[54] J.D. Herbsleb and D.R. Goldenson, "A Systematic Survey of CMM Experience and Results," *Proc. 18th Int'l Conf. Software Eng. (ICSE-18),* pp. 323-330, 1996.

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

[56] W.S. Humphrey, *Managing Technical People: Innovation, Teamwork, and the Software Process.* Reading, Mass.: Addison-Wesley, 1997.

[57] W.S. Humphrey, T. Snyder, and R. Willis, "Software Process Improvement at Hughes Aircraft," *IEEE Software,* vol. 8, no. 4, pp. 11-23, 1991.

[58] *Juran's Quality Handbook,* J.M. Juran and A.B. Godfrey, eds., fifth ed., New York: McGraw-Hill, 1999.

[59] H.F. Kaiser, "A Second Generation Little Jiffy," *Psychometrika,* vol. 35, pp. 401-417, 1970.

[60] R.S. Kaplan and D.P. Norton, *The Balanced Scorecard: Translating Strategy into Action.* Boston, Mass.: Harvard Business School Press, 1996.

[61] B.A. Kitchenham, *Software Metrics: Measurement for Software Process Improvement.* Oxford, England: NCC Blackwell, 1996.

[62] S. Konishi, "Normalizing Transformations of some Statistics in Multivariate Analysis," *Biometrika,* vol. 68, no. 3, pp. 647-651, 1981.

[63] J. Lave and E. Wenger, *Situated Learning: Legitimate Peripheral Participation.* Cambridge: Cambridge Univ. Press, 1991.

[64] J.G. March, "Exploration and Exploitation in Organizational Learning," *Organization Science,* vol. 2, no. 1, pp. 71-87, 1991.

[65] J.G. March, *The Pursuit of Organizational Intelligence.* Malden, Mass.: Blackwell, 1999.

[66] J.G. March and R.I. Sutton, "Organizational Performance as a Dependent Variable," *Organization Science,* vol. 8, no. 6, pp. 698-706, 1997.

[67] J. Mart, *Cultures in Organizations: Three Perspectives.* New York: Oxford Univ. Press, 1992.

[68] E. Mayo, *The Human Problems of an Industrial Civilization.* Boston: Harvard Univ. Press, 1933.

[69] E. Mayo, *The Social Problems of an Industrial Civilization.* Boston: Harvard Univ. Press, 1945.

[70] B. McFeeley, *IDEAL: A User's Guide for Software Process Improvement,* Handbook, CMU/SEI-96-HB-01, Carnegie Mellon Univ., Software Eng. Inst., 1996.

[71] T. Mehner, "Siemens Process Assessment Approach," *Better Software Practice for Business Benefit: Principles and Experience,* R. Messnarz and C. Tully, eds., CS Press, pp. 199-212, 1999.

[72] R. Messnarz, "Summary and Outlook," *Better Software Practice for Business Benefit: Principles and Experience,* R. Messnarz and C. Tully, eds., CS Press, pp. 389-393, 1999.

[73] D.A. Nadler, *Feedback and Organization Development: Using Data-Based Methods.* Reading, Mass.: Addison-Wesley, 1977.

[74] I. Nonaka, "A Dynamic Theory of Organizational Knowledge Creation," *Organization Science,* vol. 5, no. 1, pp. 14-37, 1994.

[75] I. Nonaka and H. Takeuchi, *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation.* New York: Oxford Univ. Press, 1995.

[76] J.C. Nunnally and I.A. Bernste, *Psychometric Theory,* third ed. New York: McGraw Hill, 1994.

[77] T.C. Powell, "Total Quality Management as Competitive Advantage: A Review and Empirical Study," *Strategic Management J.,* vol. 16, no. 1, pp. 15-37, 1995.

[78] K. Pulford, A. Kuntzmann-Combelles, and S. Shirlaw, *A Quantitative Approach to Software Management: The AMI Handbook.* Wokingham, England: Addison-Wesley, 1996.

[79] A. Rainer and T. Hall, "Key Success Factors for Implementing Software Process Improvement: A Maturity-Based Analysis," *J. Systems and Software,* vol. 62, no. 2, pp. 71-84, 2002.

[80] S. Rifkin, "What Makes Measuring Software So Hard?" *IEEE Software,* vol. 18, no. 3, pp. 41-45, 2001.

[81] J. Rooijmans, H. Aerts, and M. van Genuchten, "Software Quality in Consumer Electronics Products," *IEEE Software,* vol. 13, no. 1, pp. 55-64, 1996.

[82] J.V. Saraph, P.G. Benson, and R.G. Schroeder, "An Instrument for Measuring the Critical Factors of Quality Management," *Decision Sciences,* vol. 20, no. 4, pp. 810-829, 1989.

[83] E.H. Schein, "Culture: The Missing Concept in Organization Studies," *Administrative Science Quarterly,* vol. 41, pp. 229-240, 1996.

[84] P.M. Senge, *The Fifth Discipline: The Art and Practice of the Learning Organization.* New York: Doubleday, 1990.

[85] H.A. Simon, "Bounded Rationality and Organizational Learning," *Organization Science,* vol. 2, no. 1, pp. 125-134, 1991.

[86] D. Stelzer and W. Mellis, "Success Factors of Organizational Change in Software Process Improvement," *Software Process —Improvement and Practice,* vol. 4, no. 4, pp. 227-250, 1998.

[87] D. Stelzer, W. Mellis, and G. Herzwurm, "Software Process Improvement via ISO 9000? Results of Two Surveys among European Software Houses," *Proc. 29th Hawaii Int'l Conf. Systems Sciences,* Jan. 1996.

[88] J. Stevens, *Applied Multivariate Statistics for the Social Sciences,* fourth ed. Mahwah, N.J.: Lawrence Erlbaum, 2002.

[89] D.W. Straub, "Validating Instruments in MIS Research," *MIS Quarterly*, vol. 13, no. 2, pp. 147-169, 1989.

[90] F.W. Taylor, *The Principles of Scientific Management.* Newton Library Harper & Row, 1911.

[91] E. Thorsrud, B. Sørensen, and B. Gustavsen, "Sociotechnical Approach to Industrial Democracy in Norway," *Handbook of Work Organization and Society*, R. Dubin, ed., Chicago: Rand McNally, pp. 648-687, 1976.

[92] E. Trist, "The Evolution of Socio-Technical Systems: A Conceptual Framework and an Action Research Program," *Occasional Papers No. 2*, Toronto, Ontario: Ontario Quality of Working Life Center, 1981.

[93] R. van Solingen and E. Berghout, *The Goal/Question/Metric Method: A Practical Guide for Quality Improvement of Software Development.* London: McGraw-Hill, 1999.

[94] P.F. Velleman and L. Wilkinson, "Nominal, Ordinal, Interval, and Ratio Typologies are Misleading," *The Am. Statistician*, vol. 47, no. 1, pp. 65-72, 1993.

[95] M. Weber, C. Wright Mills, and H.H. Gerth, *From Max Weber: Essays in Sociology.* Oxford: Oxford Univ. Press, 1958.

[96] E. Wenger, *Communities of Practice: Learning, Meaning, and Identity.* Cambridge: Cambridge Univ. Press, 1998.

[97] S.M. Yusof and E. Aspinwall, "Critical Success Factors for Total Quality Management Implementation in Small and Medium Enterprises," *Total Quality Management*, vol. 10, nos. 4-5, pp. 803-809, 1999.

[98] S. Zahran, *Software Process Improvement: Practical Guidelines for Business Success.* Harlow, England: Addison-Wesley, 1998.

Tore Dybå received the MSc degree in electrical engineering and computer science from the Norwegian Institute of Technology (NTH) in 1986 and the PhD degree in computer and information science from the Norwegian University of Science and Technology (NTNU) in 2001. He is the chief scientist at SINTEF (the Foundation for Scientific and Industrial Research at the Norwegian Institute of Technology) and a visiting scientist at the SIMULA Research Laboratory. Dr. Dybå worked as a consultant both in Norway and Saudi Arabia before he joined SINTEF in 1994. His research interests include empirical and evidence-based software engineering, software process improvement, knowledge management, and organizational learning. Dr. Dybå is the author and coauthor of several publications appearing in international journals, books, and conference proceedings in the fields of software engineering and knowledge management. He is the principal author of the book *Process Improvement in Practice: A Handbook for IT Companies*, published as part of the Kluwer International Series in Software Engineering. He is a member of the International Software Engineering Research Network (ISERN) and the IEEE Computer Society.