Reasons for Software Effort Estimation Error: Impact of Respondent Role, Information Collection Approach, and Data Analysis Method

Magne Jørgensen and Kjetil Moløkken-Østvold

Abstract—This study aims to improve analyses of why errors occur in software effort estimation. Within one software development company, we collected information about estimation errors through: 1) interviews with employees in different roles who are responsible for estimation, 2) estimation experience reports from 68 completed projects, and 3) statistical analysis of relations between characteristics of the 68 completed projects and estimation error. We found that the role of the respondents, the data collection approach, and the type of analysis had an important impact on the reasons given for estimation error. We found, for example, a strong tendency to perceive factors outside the respondents' own control as important reasons for inaccurate estimates. Reasons given for accurate estimates, on the other hand, typically cited factors that were within the respondents' own control and were determined by the estimators' skill or experience. This bias in types of reason means that the collection only of project managers' viewpoints will not yield balanced models of reasons for estimation error. Unfortunately, previous studies on reasons for estimation error have tended to collect information from project managers only. We recommend that software companies combine estimation error information from in-depth interviews with stakeholders in all relevant roles, estimation experience reports, and results from statistical analyses of project characteristics.

Index Terms—Cost estimation, review and evaluation, performance evaluation.

1 Introduction

In [1], we found, through a review of surveys on software estimation, that the average effort overrun of software projects seems to be in the range 30 to 40 percent,[1] i.e., the average estimation error of software projects is high. In order to reduce the estimation errors, we need to have means of understanding *why* estimation errors occur. Important questions for that purpose are: Should the software organization base their collection of error information on interviews, project reviews, or statistical analyses of project characteristics? What is the relation between the types of reasons for error provided and the data collection approach? Previous studies on reasons for estimation error (see Section 2) have been based mainly on questionnaires to project managers and statistical analysis of project characteristics. Does this bias our understanding of why estimation errors occur? This paper aims at answering these questions and is based on a study conducted within one medium-large (about 100 employees) Norwegian software development organization. Our goal is to identify how different roles, information collection approaches, and analysis techniques may supplement each other and lead to better, and more comprehensive models of why estimation errors occur, i.e., our focus is *not* the identification of the most important reasons given for estimation error in the studied organization, but the *processes* by which a better understanding of why estimation errors occur may be gained.

The main reason for our focus on the *process* of understanding why estimation errors occur, rather than on the estimation errors themselves, is that we believe that companies should attempt to understand why estimation errors occur in their own particular context, and that it may be difficult to learn much from general studies on estimation errors in other companies. For example, assume that we ask project managers A, B, and C in three companies about the most important reason for estimation errors. Project managers A and B cite "overlooked tasks" and project manager C "immature users" as the most important reasons. To conclude from this, finding that "overlooked tasks" is more important than "immature users" may not be a very useful conclusion for companies similar to the company of project manager C.

This leads to the *research questions* of our study:

- Are there differences in the reported types of reasons for estimation error which are dependent on the organizational role of the respondent, e.g., whether the respondent is a project manager or a general manager?
- Are there differences in the reported types of reasons for estimation error which are dependent on the

1. We are aware of the study by Standish Group which reports an average 189 percent overrun. That study, however, seems to have methodological weaknesses. See our discussion in "*How large are software cost overruns? Critical comments on the Standish Group's CHAOS Reports*" (preliminary version can be downloaded from http://www.simula.no/publication_one.php?publication_id=711).

- *The authors are with Simula Research Laboratory, PO Box 134, N-1325 Lysaker, Norway. E-mail: {magne.jorgensen, kjetilmo}@simula.no.*

Manuscript received 2 Dec. 2003; revised 27 Oct. 2004; accepted 3 Nov. 2004. Recommended for acceptance by M. Shepperd.
For information on obtaining reprints of this article, please send e-mail to: tse@computer.org, and reference IEEECS Log Number TSE-0196-1203.

information collection method, e.g., whether interviews or project experience reports are used?

- Are there differences in the reported types of reasons for estimation error which are dependent on the analysis method, i.e., whether applying a qualitative analysis of project experience reports or a quantitative (statistical regression) analysis of the same project data?

The remaining part of the paper tries to answer these questions and is organized as follows: Section 2 describes previous work related to the research questions. Section 3 describes the design of the study, including limitations and challenges. Section 4 reports and discusses the results of the study. Section 5 concludes and describes further work.

2 Related Work

Most published studies on reasons for software development estimation errors are based on questionnaires. The design and results of these questionnaire-based studies are briefly described in Table 1. In addition, a few studies on estimation errors have been based on statistical analyses. Table 2 summarizes those studies.

There are clear similarities in the results in the studies in Table 1. Most studies seem to focus on immediate (direct) reasons for estimation error and reasons related to the clients and users, e.g., "frequent changes," "frequent requests for changes by users," "changing requirements and specifications," and "requirement change/addition/deletion." The studies in Table 1 have a predominance of project managers as respondents, and simple questionnaires as the approach to data collection. This predominance of project managers as respondents may have caused the focus on factors outside the project managers' own control when providing reasons for estimation errors. None of the estimation error reason questionnaires stimulated to the provision of comprehensive reasoning models. The emphasis of direct reasons is, however, not mainly due to the predefined reasons of the questionnaires since most of the predefined reasons were indirect, i.e., reasons not directly connected to estimation error.

The statistical analysis-based results described in Table 2 are different from the questionnaire-based reasons found in Table 1. The difference in reasons is probably not only a result of difference in the variables collected and analyzed, but may also have been caused by differences in the method of analysis. For example, while the statistical analyses suggest that the size of the project is an important indicator of estimation error, the interviews, estimation experience reports, and questionnaires did not mention size of project as a factor affecting estimation error. This may mean that there are relationships that are easier to examine through statistical analysis than through "on the job" experience. In other words, the studies in Tables 1 and 2 suggest that experience-based reasons and associations found through statistical analyses supplement, rather than replace, each other.

Table 3 summarizes findings from two interesting studies on reasons for estimation error in manufacturing and construction projects.

These two studies demonstrate, perhaps even more clearly than the software studies, the importance of the respondents' role when providing reasons for errors and failures. This finding is supported by Tan and Lipe [10]. They found, in a business management context, that low estimation accuracy was explained as due to uncontrollable external factors.

3 Design of Study

The studies reported in Section 2 were not designed for the purpose of investigating differences in types of reasons for estimation error dependent on the respondents' role, information collection method, and analysis technique. The studies have, for example, not separated estimation error reasons reported by project managers from those reported by software developers or organizational managers. They have not analyzed how different data collection methods may lead to an emphasis on different estimation error reasons within the same organization. This section describes a study better designed for this purpose.

3.1 Types of Reasons

What does it mean to believe that an event or characteristic, e.g., a major unexpected change in requirements during project execution, is a reason for effort overruns in software projects? The answer to this question is not trivial. Potential interpretations of something (X) being a reason for effort overruns are, for example:

- There is a *direct* causal link between X and the overrun, i.e., X is a *direct reason* for overrun.
- X leads to events that, in turn, lead to overruns, i.e., X is an *indirect reason* for overruns. If the events leading to overrun started with X, we may call X the *root reason* or the *trigger reason*.
- The events actually leading to overrun would have been harmless if X had not been present, i.e., X is an important *contributory reason*, or *necessary condition* for the overruns.
- The overrun increases when X is present, i.e., X is a *deterministic* reason.
- The presence of X increases the probability of overrun, i.e., X is a *probabilistic* reason.
- Mainly the large overruns were caused by X, i.e., X is mainly a *large overruns* reason.
- The main contributor to high *average overrun* is X, i.e., X is an *average-overruns* reason.

More discussions on types, definitions, and interpretations of reasoning models can be found in [11].

For the purpose of this study, we decided to focus on direct reasons, indirect reasons, and contributory reasons. Our interpretation of these types of reasons is illustrated below:

- *Direct reason*. A reason is categorized as a direct reason if the estimation error is explained by an immediate reason for estimation error. For example, "unexpected problems with the testing tool" is a reason that may immediately lead to estimation error.

TABLE 1
Questionnaire-Based Studies on Reasons for Software Estimation Error

Study	Population	Study Design	Results
Phan et al. [2]	Software professionals (80% of them were project managers or developers) in 191 organizations.	Four pre-defined categories: Long duration, over-optimism, poor analysis and design, and frequent changes.	The two most important reasons were "unrealistic over-optimism[2]" and "frequent changes".
Van Genuchten [3]	Project managers responsible for the estimation of 160 activities in six development projects within one department.	Pre-defined classification of reasons for error. The six project managers marked one (or more) of these for each activity.	Most frequent reasons were "more time spent on other work than planned" and "complexity of application underestimated".
Lederer and Prasad [4]	Estimation responsible (mainly project managers and developers) personnel in 112 organizations.	Pre-defined list of reasons where general importance for estimation error was marked with a value from 1 to 5.	Most important reasons were "frequent requests for changes by users", "users lack of understanding of their own requirements", and "overlooked tasks".
Standish Group - 1994[3]	"IT executive managers" (mainly project managers?) from 365 organizations.	Pre-defined classification of reasons.	The three most important reasons for estimation over-runs were "lack of user input", "incomplete requirements and specifications", and, "changing requirements and specifications".
Subramanian and Breslawski [5]	Project managers in different companies representing 45 projects.	Reasons classified by the authors based on responses from the project managers.	Most important reasons were "requirement change/addition/deletion", "programmer or team member experience, turnover", and, "design changes, scope, complexity".

2. It is unclear how to interpret "unrealistic overoptimism" as a reason for estimation overruns. To some extent, "unrealistic overoptimism" is the same thing as "effort overrun."
3. www.standishgroup.com/sample_research/chaos_1994_1.php. There are commercially available updates of the 1994 report available.

- *Indirect reason*. A reason is categorized as an indirect reason if the estimation error is explained by reasons not directly connected to estimation error. For example, "lack of project manager culture" may lead to "insufficient effort on project planning," which in turn may lead to "overlooked tasks" and estimation error. "Lack of project manager culture" and "insufficient effort on project planning" are both indirect reasons of different distance to the direct reason "overlooked tasks." This category also covers more complex mental models of reasons, such as multiple reasons with joint effect on estimation error.
- *Contributory reasons*. A reason is categorized as a contributory reason if the reason is better described as a necessary condition of estimation error than a direct or indirect reason. For example, assume that "overlooked tasks" is considered to be the direct

TABLE 2
Statistical Analyses of Factors Associated with Estimation Error

Study	Population	Design of Study	Results
Lederer and Prasad [6]	Same dataset as in [4], see Table 1.	Analysis of statistical significance of differences.	Significant ($p<0.05$) findings were that the estimation error increased with use of estimation tools, when not estimating own work, when there was no revision of estimates by the management, when there was no independent evaluation of development process, when there was no formal process of cost control, and, when there was no evaluation of estimation accuracy to assess the managers, the estimators or developers.
Standish Group, Chaos Report, 1994	"IT executive managers" (mainly project managers?) from 365 organizations.	Analysis of difference in mean effort overrun of different types of project.	Estimation error increased with increased size.
Gray et al. [7]	Information about the development of 77 modules of a large health-care system.	Several types of statistical analysis associations between estimation error category and module properties.	The analyses showed, amongst other things, that over-estimation was connected with changes on small modules, development of screens and modules accessing one or less data tables, while under-estimation was connected with changes on large modules, development of reports, and models accessing more than one data table.

reason of estimation overrun in a project. A contributory reason of the estimation overrun could then be "lack of estimation checklists," i.e., "overlooked tasks" could have been prevented by better estimation checklists.

We focus on these three categories of reason because the categories enable a separation between simple, complex, and condition-dependent types of reasoning models. We believe that there may be important relationships to be identified based on these categories. For example, Brickman et al. [12] found when studying car accidents, that the involved drivers dominantly reported direct reasons, while external observers reported more indirect reasons.

3.2 Data Collection

Company: The company that provided the data is a medium-large (about 100 employees) Norwegian software development company that produces Web-portals, e-commerce solutions and content management systems for their clients. The main work process is based on the waterfall development model and contains six phases: strategy and concept, specification, development, test, implementation, and evaluation. There was some use of evolutionary-incremental development models. Most projects were "multidisciplinary," i.e., they involved professionals in the role of "graphic designer," "user interaction designer," "project manager," and "software developer." The company had not implemented a formal estimation process and the actual estimation processes, consequently, varied within the organization. The dominant estimation approach was bottom-up, i.e., work breakdown structure-based. Usually, the project manager was responsible for the project's estimate. Most projects were small and there were many different clients, whose experience level regarding software projects varied from those who requested their first Web system to companies that had based their daily operations on software systems for many years. Most of the estimates were completed as part of the project planning or bidding process.

The main organizational roles typically involved in the estimation of effort were 1) general managers with responsibility for the bidding and contracting processes, 2) project managers with responsibility for the project's total effort estimates, and 3) software developers with responsibilities for the effort estimation of design, programming, and testing activities of the project. In some cases, e.g., when the

TABLE 3
Two Manufacturing and Construction Projects

Study	Population	Design of Study	Results
Thambain and Wilemon [8]	304 participants in project management workshops and seminars.	Pre-defined categories on reasons for cost and time overrun. Use of questionnaires.	The general managers perceived that "insufficient front-end planning", and, "unrealistic project plans" were the most important reasons, while the project managers believed that "client/management changes" and "technical complexity" were the most important reasons.
Chan and Kumaraswamy [9]	147 organizations involved in construction projects in the role of clients, consultants, or contractors.	Pre-defined categories of reasons for delays in building and civil engineering projects. Use of questionnaires.	In building projects the clients believed that "poor site management and supervision", and, "inadequate managerial skill" were the two most important reasons for delays. The consultants also believed that "poor site management and supervision" was the most important reason, but included "unforeseen ground conditions" as the second most important reason. The contractors, i.e., the organizations responsible for the delay, believed that the delays were mainly caused by "delays in design information", and, "long waiting time for approval of drawings".

projects were small, software developers could be responsible for the effort estimates of the whole project. For larger projects, the project leader would typically arrange a meeting with experienced software developers and derive an effort estimate of the project through group estimation.

3.2.1 Limitations Related to Selection of Company and Size of Projects

The studied company had mainly small projects, informal estimation and development processes, and, dominantly, immature clients. This means that other types of company may give different reasons for estimation error. The *types* of reason provided and the impact from data collection and analysis approach to the types of reason may, however, be more robust toward size of projects, formality of processes, and maturity of clients. For example, the reasons for estimation error may be different for small projects and large projects. The impact of the role of the respondent on the *type of reasons* provided, on the other hand, is, as far as we have observed from discussion with managers of projects of different sizes, much less impacted by the size of the project.

3.2.2 Data Collection Approaches

To examine the impact of the role of respondent, data collection approach, and analysis technique, we decided to collect reasons for estimation error based on three approaches in the same organization, i.e., 1) through general *interviews* with eight employees responsible for the estimates, 2) through 68 *project estimation experience reports*, and 3) through *statistical analysis* of associations between project characteristics and estimation error of the same 68 projects as in 2). None of the interviewed employees provided project experience reports. The 68 project experience reports were completed by 29 different employees in the roles of project manager or developer. The data collection lasted about one year.

3.2.3 Limitations of the Data Collection

The interviews and the projects do not describe exactly the same estimation situations because they were collected at different times and because we excluded projects with planned duration of more than four months from the logging. Both limitations were the result of practical concerns. The permission to log information about projects came as a result of the analysis of the interviews. We initially intended that the logging would not last more than a few months. Projects longer than four months would consequently be difficult to complete within our study. There were, however, no large changes in estimation or development process in the period between the interviews and the project logging, and the company had very few large projects. Nevertheless, the limitation may have had an impact on the difference in the reasons for estimation error provided in the interviews, and in the project data. We do not believe, however, that this limitation has an important impact on *how* estimation error reasons are described, i.e., *what types* of reasons people give to explain estimation errors. As stated earlier, our goal is not to examine the reasons for estimation errors in the studied company, but to analyze the impact of the role, the data collection approach, and the analysis approach on the types of error reasons given.

3.2.4 Interviews

One of the authors of this paper interviewed the following eight management personnel responsible for estimates:

- The manager of the technical personnel (M-Tech).
- The manager of the human-computer-interaction personnel (M-HCI).
- The manager of the graphic design personnel (M-Graph).
- The most senior project manager (PM-Sen). This project manager was frequently used to review other project managers' estimates.
- Two project managers with technical background (PM-Tech1 and PM-Tech2).
- A project manager with human computer interaction background (PM-HCI).
- A project manager with graphic design background (PM-Graph).

Following an introduction about the purpose of the interview and general questions about the estimation process, we asked the above-mentioned personnel to give reasons for both accurate and inaccurate effort estimates. No predefined categories of reason or templates for the answers were used. We instructed the interviewed employees to base their reasons on experience from a large set of projects, and not, for example, one or two projects with especially large overruns. Each interview lasted 1-2 hours. The interviews were meant to be the organization's first step towards the improvement of the estimation process. However, due to reorganizations and other unexpected events, the planned estimation improvement work was never continued.

3.2.5 Experience Reports and Statistical Analysis of Project Characteristics

Over a period of approximately a year, we collected information about projects with an estimated effort of less than approximately four calendar months, i.e., we excluded the largest projects. In total, information about 68 projects was collected. The effort used in these projects (or tasks) varied from four work-hours to 1,683 work-hours, with a median of 45 work-hours. All these 68 projects provided "estimation experience reports," where reasons for accurate or inaccurate estimates were provided, together with other information about the project. The chief project manager of the company was in charge of data collection. He asked the estimators to complete one questionnaire just after completing the project planning and another one after the project was completed. The completion of the questionnaires was supported by a spreadsheet-based tool that guided the project manager through a number of questions.

The information collected before a project started was as follows (with predefined categories):

- Company role of the estimator (Project manager, developer).
- Brief description of the project (Free text).
- The estimators' assessment of the complexity of the project (Easy, medium, complex).
- Type of contract (Payment per hour, fixed price).
- The estimators' assessment of how important the project is for the client (Low/medium importance, high importance, critical).
- The priority that the client assigns to the project (Cost, quality, or time-of-delivery).
- The estimators' self-assessed level of knowledge about how to perform the project (Know something about how to solve the task, know a great deal about how to solve the task).
- The estimators' planned participation in the completion of the project (0, 1-50, 51-100 percent of the work planned to be completed by the estimator him/herself).
- The estimators' perception of his typical accuracy when estimating similar projects (accuracy categories from "less than 10 percent" to "more than 100 percent").
- The estimated effort in work hours. (We found that the estimators had slightly different interpretations of "estimated effort." In most projects, however, "estimated effort" seemed to be interpreted as "planned effort," i.e., "most likely effort" added a contingency buffer. All remaining project activities were included in the estimate, e.g., project administration, design, programming, and test.)

After the project was completed the estimators provided an estimation experience report in terms of:

- The actual effort in work hours.
- Comments on the actual use of effort.
- Descriptions of unexpected problems during the execution of the project.
- Reasons for high or low estimation accuracy.

All project characteristics were included in the statistical analyses of estimation error. The estimation experience report was based on all information collected immediately after the completion of the project, in particular, the responses when we asked for "reasons for high or low estimation accuracy."

3.3 Measures

We apply two common measures of estimation error in this study. One measure is of the mean magnitude of relative error (mean MRE) and the other is of the mean relative error (mean RE). The mean relative error shows the *bias* of the estimates, e.g., a high RE means that there is a strong tendency to underestimation.

Mean MRE (Estimation Error) is measured as:

$$\frac{1}{n}\sum_i \frac{|Act_i - Est_i|}{Act_i},$$

where Act_i is the actual effort on project i, Est_i is the estimated effort for project i, and n is the total number of projects.

Mean RE (Bias) is measured as:

$$\frac{1}{n}\sum_i \frac{(Act_i - Est_i)}{Act_i}.$$

In situations with both effort overruns and underruns, the mean RE (bias) will have a lower absolute value than the mean MRE (estimation accuracy). The reason for this is that, when calculating the bias, similar degrees of overruns and underruns cancel each other. Assume, for example, two projects with actual efforts of 100 work-hours. The first project was estimated to 75 work-hours and the other to 125 work-hours. The mean estimation accuracy (mean MRE) of these two projects is 25 percent, while the mean bias (mean RE) is 0 percent.

4 Results

4.1 Interviews

Table 4 describes the most important reasons (or, in some cases, reasoning models) for estimation error as perceived by each interviewed subject according to our categorization of, and notation for, reasons (direct reason $\rightarrow$, indirect reason $\rightarrow\rightarrow$, or contributory reason $\downarrow\rightarrow$). We have translated and condensed the most important reasons provided by each subject without, as far as we are aware, changing the intended opinion of the respondents. To evaluate interrater reliability the two authors and an experienced software professional independently analyzed the transcripts of the interviews. A comparison of the reasons identified by the analysts revealed no serious conflicts in the interpretation of the reasons provided by the interviewed subjects. There was, however, some differences in what the analysts believed were perceived as the most important reasons by the interviewed subjects and the amount of reasons identified. We decided to include a reason if it was identified by at least one of the analysts.

There are several interesting observations that can be derived from the interviews summarized in Table 4:

- Although there are common patterns in the responses, e.g., the need for better learning opportunities, the role of the respondent seems to have a strong bearing on the type of reasons provided. For example, there seems to be a pattern that general managers (M-Tech, M-HCI, M-Graph) more frequently provide more general reasons for estimation error than the project managers (PM-Sen, PM-Tech, PM-HCI, PM-Graph). In addition, there seems to be a tendency *not* to criticize work connected to one's own role, e.g., few of the project managers pointed at poor project planning or management as important reasons for estimation error. In other words, the factors cited for estimation error have a tendency to be outside the control of the respondent. This pattern of not criticizing factors controlled by oneself is consistent with the results of several of the studies reported in Section 2.
- Only one of the respondents provided reasons that were described as contributory reasons ($\downarrow\rightarrow$), i.e., important enablers of estimation error outside the main chain of reasons leading to estimation error. We obviously need more observations to evaluate whether this is typical or not, but there may be a need for explicit focus on contributory reasons when these are important.
- Frequently, the steps in the chain from an indirect reason ($\rightarrow\rightarrow$) to the estimation error were not well explained. For example, PM-Sen claimed that "insufficient standardization of planning and development processes" is a reason for estimation error. More standardization is, however, no "silver bullet" when improving estimation accuracy. Its impact on estimation accuracy depends, among other things, on properties of the standards applied, and the organization's ability to establish processes that enable learning from experience with standardized processes. To really understand the provided reasons for estimation error, we may have to push the respondents for more comprehensive structures of reasons, where all important steps and all non-obvious contributory reasons are included.
- All of the reasons were described deterministically, none probabilistically. This suggests that the models cited by the respondents to explain estimation overruns are deterministic and not probabilistic. Hammond [13] suggests that the ability to understand relationships in terms of probabilities instead of purely deterministic connections is important for correct learning in situations with high uncertainty, such as effort estimation of software projects. For example, instead of the deterministically described reason for estimation errors: "*Clients unable to deliver a good requirement specification, leads to unplanned rework*." (PM-Tech1), a probabilistic description of the same reality may be: "When clients are not able to deliver a good requirement specification, this leads to a higher probability of unplanned rework." The latter description emphasize that the lack of a good requirement specification does not always lead to inaccurate estimates, i.e., that there is a probabilistic relationship between quality of requirement specification and estimation error. The ability to think about and describe reasons in probabilistic terms can, according to Brehmer [14], hardly be derived from experience alone, but must be taught. We have included a more comprehensive discussion about the importance of probability-based reasoning models when learning from software estimation experience in [15].
- It was frequently unclear whether the respondents described reasons for the largest overruns or the typical overruns, i.e., the scope of the reasons were not described.

TABLE 4
Interview-Based Reasons for Estimation Error

Subject	Reasons
M-Tech *(Manager of the software developers)*	No systematic feedback to enable learning (→→). Insufficient time on estimation and planning (→→), leads to overlooked tasks (→).
M-HCI *(Manager of the HCI personnel)*	Lack of processes enabling learning from experience (→→). Insufficient focus on HCI in the estimation process (→→). Lack of client realism in HCI-requirements (→→). Poor project planning (→→). Poor project management (→→).
M-Graph *(Manager of the graphical designer personnel)*	Project managers are not skilled in planning multi-disciplinary projects (→→), which leads to insufficient focus on graphic design in the estimation process (→→), and inefficient allocation and use of graphic design resources (→). No systematic feedback to enable learning (→→). Insufficient tool support for project management (→→). Poor project management (→→). Customer requirements difficult to interpret (→→).
PM-Sen *(Senior project manager with extensive experience from project bidding and planning)*	Insufficient focus on the project manager role (→→), leads to insufficient training and feedback (→→). Insufficient standardization of planning and development processes (→→). The experience database of previous projects is not used (→→). Inefficient allocation of project resources (→→).
PM-Tech1 *(Project manager with technical background)*	Clients unable to deliver a good requirement specification (→→), leads to unplanned re-work (→). Lack of requirement change control processes (→→). Insufficient time spent on estimation and planning (→→). Not sufficient focus on learning from experience (→→).
PM-Tech2 *(Project manager with technical background)*	Projects are frequently different from earlier projects (→→), leads to lack of relevant experience when estimating (→), because of lack of checklists (→) and experience database (→). Incomplete requirement specifications (→→).
PM-HCI *(Project manager with HCI background)*	HCI is involved too late (→→), which leads to unrealistic expectations by clients (→→), and unplanned activities (→). Project manager has insufficient knowledge about HCI (→→). Not sufficient focus on learning from experience (→→).
PM-Graph[4] *(Project manager with graphic designer background)*	Insufficient focus on graphic design in the estimation process (→→). No systematic feedback to enable learning (→→). Estimate strongly impacted by price-to-win (→→). Lack of justification of estimates (→→).

4. Important comment from PM-Graph: "We seldom have large overruns on graphic design activities. A graphical design can be completed in five hours or five months, dependent on how much money the client is willing to spend to ensure a good design, for example, on iterations and user tests."

Interviews may enable the description of complex models of reasons. However, our interviews suggest that models of more than 2-3 steps, with contributory reasons, with probabilistic relationships, and of well-defined scope, are not provided by software professionals when simply asked for "reasons for estimation error," i.e., there may be a need for more structure in the process of elicitation of comprehensive models of estimation error. In addition, there may be a need for elicitation structures that support the identification of estimation error reasons that are avoided because they may be interpreted as excuses, e.g., "bad luck," and reasons that are difficult to measure, e.g., the impact from the "the winners curse" (the organization only gets the contract when the bids are based on over-optimistic effort estimates).

4.2 Project Experience Reports

Based on repeated readings of the 68 experience reports, one of the authors developed a classification of reasons for estimation errors. This classification process was based on joining reasons perceived to be closest to each others until a "reasonable number" (15 in our case) of categories had been identified. We decided not to use any a priori defined classification schema. Applying a classification tailored to the actual formulation in the experience report made our categorization simpler, e.g., less need for translation between the formulation and the category. Not using an a priori classification is not an important limitation of this study since the main purpose is not to compare the reasons of estimation error of this study with other studies, but to analyze how roles, information collection approaches, and analysis methods impact how estimation errors are explained.

TABLE 5
Experience Report-Based Reasons for Inaccurate Estimates

Id.	Reason	Reported in Project	Mean MRE	Mean RE	Proportion of Over Median Large Projects
1	Unexpected events and overlooked tasks (→)	5, 8, 10, 11, 15, 21, 25, 26, 30, 31, 35, 43, 47, 49, 50, 51, 52, 58, 60, 61, 62, 63, 64, 65, and, 66	0.32	0.32	60%
2	Change requests from clients or "functionality creep" (→)	5, 7, 9, 14, 15, 16, 18, 22, 23, 31, 47, 48, 61, and, 67	0.35	0.32	71%
3	Simpler task or more skilled developer than expected (→)	13, 34, 36, 42, 57, and, 59	0.54	-0.54	17%
4	Resource allocation problem (→→)	8, 28, 43, and, 47	0.32	0.32	50%
5	Poor requirement specification or problems with communication with the client (→→)	4, 8, 18, 22, 25, 26, 31, 43, 44, 45, 48, 54, 59, 63, and, 67	0.42	-0.26	73%
6	Too little effort on estimation work (→→)	63	0.70	0.70	100%
7	High priority on quality, cost accuracy not of high importance (→→)	17, 18, 22, and, 30	0.32	0.33	75%
8	More reuse than expected from other projects (→)	4, and, 57	0.61	-0.61	50%

Three independent raters, the two authors and an experienced software professional, were applied for the purpose of classifying the estimation error reasons described in the experience reports. The results from the three raters suggested that the classification schema was appropriate, i.e., there was no strong need for additional classes of reasons or reformulation of existing categories. An analysis of differences between the raters showed, however, the usefulness of having three independent raters. In 60 percent of the experience reports, at least one rater identified an additional reason not found by the other two raters. We investigated carefully the reasons not identified by all raters and found in most cases that these reasons could be defended from the text in the experience reports. In several cases, the same explanation in the experience reports could be interpreted differently, e.g., the same formulation could be interpreted as "functionality creep" or as "overlooked task." In such cases, we decided to include both types of reasons.

Table 5 (types of reasons for inaccurate estimates) and Table 6 (types of reasons for accurate estimates) summarize the reasons provided in the experience reports. The estimators themselves decided whether they considered the estimate to be accurate or inaccurate, i.e., whether they should report reasons for accurate or inaccurate estimates. For each set of projects belonging to a particular category of reason, we calculated the mean MRE (estimation accuracy), the mean RE (bias), and the proportion of "over median large projects." The median effort of the studied projects was 45 work-hours, and all projects with estimated effort more than 45 work-hours were hence classified as "over median large." Most projects mentioned more than one reason for estimation inaccuracy or accuracy and several experience reports described reasons for both accuracy and inaccuracy. There were typically good reasons for including

TABLE 6
Experience Report-Based Reasons for Accurate Estimates

Id.	Reason	Reported in Project	Mean MRE	Mean RE	Over Median Large Projects
1	Inclusion of a large buffer to deal with unexpected events and/or changes in specification (→)	7, 9, 25, 29, 32, 45, 55, and, 59	0.18	-0.10	25%
2	Simple project (→)	1, 2, 6, 13, 15, 19, 20, 24, 27, 37, 53, 55, and, 56	0.12	-0.03	23%
3	Experience from a similar project (→→)	1, 6, 9, 10, 12, 19, 21, 28, 33, 41, 51, and, 55	0.09	-0.01	42%
4	Good knowledge of how to solve the requirement specification (→→)	3, 9, 12, 14, 16, 21, 28, 29, 33, 37, 46, 53, and, 68	0.20	-0.14	54%
5	A high degree of flexibility in how to implement the requirement specification (→)	3, and, 42	0.21	-0.21	100%
6	Much time was spent on estimation work (→→)	9, 15, and 29	0.16	-0.06	33%
7	Good cost control (→→)	16, 20, 24, and, 42	0.13	-0.06	100%

both types of reason. For example, Project 7 reported reasons for both inaccuracy, i.e., "unexpected change requests," and accuracy, i.e., "large contingency buffer." The total estimation overrun of Project 7 was 11 percent. The project manager's explanation for including both reasons was that the unexpected change requests did lead to more work than planned, i.e., to inaccuracy, but the large contingency buffer saved the project's estimate and led to an overall acceptable level of accuracy.

The mean MRE of all tasks was 28 percent and the mean RE was 8 percent, i.e., the average estimation error was 28 percent and the average bias was 8 percent (underestimation). As opposed to earlier studies, see, for example, [16], we found that larger projects did not have larger estimation error or stronger tendency towards underestimation. An analysis of the mean MRE and RE of subclasses of the projects, i.e., the subclasses of projects with common estimation errors, did not change this relationship between project size and estimation error.

Interesting observations that can be derived from Tables 5 and 6 include:

- Most reasons were direct reasons. There were relatively few projects that described indirect reasons and none that described contributory reasons. The actual reasons provided in the estimation experience reports were only to some extent overlapping the reasons provided in the interviews. In general, it seems as if interview-based reasons focus more on process and learning issues, while estimation experience reports-based reasons focus more on specific events and specific project or estimator characteristics.
- Similar to in the interviews, the respondents had a tendency to report reasons *outside* their own control as reasons for estimation error. For example, "unexpected events and overlooked tasks" typically referred to events and tasks outside the control of the project. Interestingly, the respondents reported reasons *within* their own control or their own skill and experience, e.g., "inclusion of a large buffer" as factors contributing to accurate estimates. In other words, the perceived reasons for accurate estimates were not the opposite of the reasons for inaccurate estimates. This means that it may be important to collect both types of reasons, not only reasons for inaccurate estimates, to understand and improve software estimation accuracy.

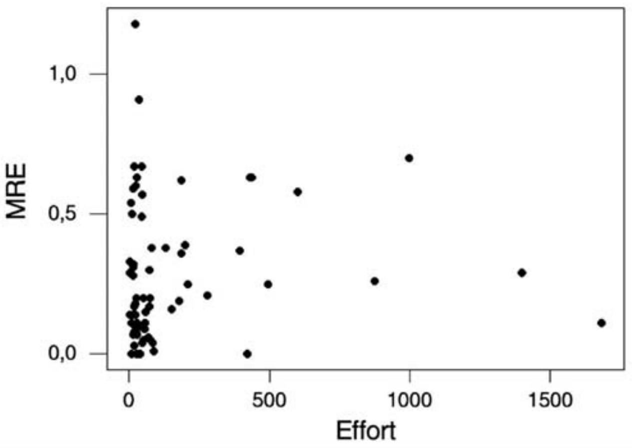

Fig. 1. Relation between effort (work-hours) and estimation error.

- The estimators regarded overestimated projects as having accurate estimates, except when the overestimation was very high. For example, the estimator of Project 42 perceived that he had delivered an accurate estimate, although he overestimated the project by 26 percent. His explanation was that the good estimation performance was a result of: "...*very strict project management, tight control of time reporting, and competent developers*." To some extent, the project manager may be correct in his interpretation of accuracy. He could easily relax the control of the project, spend more effort and, consequently, improve the estimation accuracy. From his point of view, it would have been unfair to perceive his estimate as inaccurate. A comparison of projects that provided more reasons for accurate estimates than for inaccurate estimates suggest that there was a cut-off point at about 40 percent overestimation. All projects with less than 40 percent overestimated effort provided more reasons for estimation accuracy than estimation inaccuracy. Only one project with more than 40 percent overestimation provided more reasons for estimation accuracy than for estimation inaccuracy. This means that when we ask project managers to provide reasons for estimation error we mainly get reasons for underestimation and very high overestimations, not medium-large overestimations. It may be important to be aware of this when it is important to understand reasons for medium-high overestimation, e.g., for companies in bidding situations losing contracts if the bids are unnecessarily high.
- There were no clear patterns relating reasons for estimation error to the size of the project. A possible conclusion is that the size of project does not affect the reasons for estimation error very much within the limited variation of project sizes studied in this company. Fig. 1 illustrates the lack of correlation between project effort and estimation accuracy (MRE).
- There were reasons we would expect to be reported but that few or nobody provided. One such reason is the "political estimation games" described, for example, in [17], [18]. For example, a client expects to pay a certain price for the software and the estimator is under strong pressure to reduce the initial "too high" estimate to ensure that the project can be started. We found few descriptions of such "political games" as reasons for estimation error. However, from informal meetings and lunch-discussions with some of the software developers we know that unrealistic client expectations of low cost may well have been an important factor affecting estimation overrun in many of the projects. This means that some reasons may not be mentioned because they are sensitive, or perhaps because the estimators feel uncomfortable about, for example, admitting that they sometimes succumb to pressure from clients.

Review of project-specific reasons for estimation errors, such as in the estimation experience reports studied in this study, seems to stimulate to description of direct reasons. It may, therefore, be necessary to actively stimulate the provision of indirect and contributory reasons to get a broader picture and stimulate so-called double-loop learning [19], i.e., learning that includes better understanding of the core factors that affect estimation error. In addition, to understand how "political games" affect estimation overruns, we may need structures that provide incentives for the stating of sensitive reasons.

4.3 Statistical Analysis

Earlier [20], we applied a *subset* of the dataset applied in this paper, i.e., the 49 earliest out of the current set of 68 projects, to develop a regression model for the prediction of estimation error.[5] The regression models of MRE (absolute estimation error) and RE (bias) were developed by applying stepwise regression with backward elimination and an alpha-value of 0.1 to remove variables. As described earlier, MRE is a measure of the *magnitude* of the estimation error and RE is a measure of the *bias* of the estimation error. A mean MRE of 10 percent means that the average relative estimation error is 10 percent. A mean RE of 10 percent, on the other hand, means that, on average, there is a 10 percent underestimation of effort.

The variables, i.e., all the project characteristics described in Section 3.3, were coded as binary variables. A full description of the coding and its rationales are provided in [20]. The resulting regression models were the following:

```
MRE = 0,14 + 0,13 Company Role
      + 0,13 Participation
      + 0,13 Client Priority,
  (p=0.03)  (p=0.08)  (p=0.07)  (p=0.09)
RE = 0,12 - 0,29 Company Role
     + 0,27 Previous Accuracy
  (p=0.05)  (p=0.004)  (p=0.01)
```

5. The previous paper focused on the possibilities of applying regression analysis as a means to *predict* the estimation error, i.e., it has a purpose different from the focus of processes for *understanding* factors with impact on the reasons for estimation error present in this paper. In addition, the previous paper focuses on the application of quantitative data, i.e., not on interviews and experience reports.

The variables included in the proposed models were defined as follows:

- Company Role: The project was estimated by a software developer = 1. The project was estimated by a project manager = 0.
- Participation: The estimator estimated the work of others = 1. The estimator participated in the estimated project = 0.
- Client Priority: The client prioritized time-to-delivery = 1. The client had other project priorities than time-to-delivery, i.e., cost or quality = 0.
- Previous Accuracy: The estimator believed that he/she had estimated similar tasks with an average error of 20 percent or more = 1; less than 20 percent error = 0.

The adjusted R^2-values were low, i.e., 11 percent for the MRE-model and 21 percent for the RE-model. This indicates that the models only explained small proportions of the variances of mean estimation errors.

A reanalysis of the project data, including the new projects, i.e., with 68 projects instead of 49, led to the following regression models:

```
MRE = 0,14 + 0,13 Company Role
      + 0,14 Participation
      + 0,14 Client Priority,
  (p=0.01)  (p=0.07)  (p=0.04)  (p=0.05)
RE = 0,10 - 0,22 Company Role
     + 0,23 Previous Accuracy
  (p=0.08)  (p=0.02)  (p=0.03)
```

The adjusted R^2-value was the same as before (11 percent) for the MRE-model, while it decreased for the RE-model (from 21 percent to 11 percent). As can be seen, the same variables were significant in the updated, as well as in the original models, with almost the same regression coefficients and similar p-values. This suggests that the regression-based relationships, in particular the MRE-model, are robust toward extensions of the dataset. Regression models are based on the assumption that the residuals are normally distributed, independent, with a constant variance, and mean value that equals zero. A validation of the regression models should therefore include an examination of the residuals. A visual examination of the residual plots gave that these assumptions were to a large extent met. The residuals, for MRE and RE, were close to, although not completely normally distributed. In particular, the residual for the MRE-values showed a bias toward negative residuals. This suggests that the MRE-model is, to some extent, misspecified. We tried several transformations of the model variables, but were unable to find better models. It is important to be aware of these limitations and possibility of misspecification of the models and interpret the results from the regression analysis carefully.

Table 7 shows the relations between estimation error and other project variables through a description of mean MRE and RE for all variable categories.

Examining the regression model and Table 7, we find that the statistical analysis supports and extends the understanding of the reasons found by interviews and estimation experience reports. For example:

- The statistical analyses, i.e., the regression models and the differences in mean MRE and RE, point to potential differences between project managers and software developers. Although the project managers had the most accurate estimates (MRE), they were also more biased toward underestimation (positive RE values). The MRE results support the importance of skilled project managers as reported in the interviews. The RE results, however, point to a need for further data collection and analysis to examine the reasons for the stronger overoptimism of the project managers.
- The statistical analyses point to higher uncertainty (and higher degree of overoptimism) when there is no participation in the project by the estimator, and the interviews suggest that HCI and graphic design work are not properly understood by many project estimators. Both descriptions support the hypothesis that stronger involvement in the project work leads to more accurate estimates.
- The statistical analyses point to the higher uncertainty of projects with a focus on time-to-delivery, and the interviews and the experience reports focus on the lack of good requirement specifications or frequency of unplanned changes. Combining these two information sources, we may state the hypothesis that the explanation for lower estimation accuracy in situations with priority on time-to-delivery is that too short a time is taken for the development of proper requirement specifications. This, in turn, may lead to unplanned changes and effort overruns. An alternative hypothesis, also consistent with the other information sources, is that projects with focus on time-to-delivery lead to a different project management focus, i.e., budget control has less focus compared with projects with focus on cost or quality. Here, we see that a diversity of information sources may support the building of more comprehensive reasoning models and alternative hypotheses than single information sources.

4.4 Comparison

To further demonstrate the impact of the data collection and analysis approach, Table 8 displays the three most frequently reported reasons for estimation error from the interviews, the three most frequently reported reasons from the experience report, and the three most important reasons derived from the statistical analysis. For comparison reasons, we only include the reasons for inaccurate estimates, not reasons for accurate estimates or analysis of bias.

Table 8 demonstrates that how we ask and how we analyse has an impact on the estimation error reasons considered as most important. As can be seen, the reasons are not inconsistent, they may to some extent be different formulations of the same underlying reason, and they are descriptions with different levels of granularity.

TABLE 7
Project Characteristics in Relation to Mean Estimation Error and Bias

Variable	Categories (# observations)	Mean MRE	Mean RE
Role of Estimator	Developer (22)	0.33	-0.05
	Project manager (46)	0.25	0.14
Complexity	Easy (27)	0.25	0.11
	Medium (29)	0.31	0.05
	Difficult (12)	0.26	0.09
Type of payment	Fixed price (38)	0.30	0.10
	Per hour (30)	0.25	0.06
Importance	Low/medium importance (20)	0.29	0.03
	Very important (40)	0.24	0.09
	Critical (8)	0.47	0.18
Priority	Cost (18)	0.22	0.01
	Quality (29)	0.24	0.11
	Time-to-delivery (21)	0.38	0.09
Knowledge on how to solve the project	Much knowledge (45)	0.29	0.04
	Some knowledge (23)	0.26	0.16
Estimation error of similar projects	0-10% error (21)	0.29	0.05
	11-20% error (26)	0.29	0.05
	21-30% error (7)	0.19	0.04
	31-50% error (6)	0.37	0.37
	51-75% error (1)	0.20	0.20
	> 76% error (2)	0.47	0.47
	information not provided (5)	0.17	-0.10
Participation of estimator in project	No participation (28)	0.31	0.17
	1-50% of total work (23)	0.21	-0.04
	51-100% (17)	0.32	0.06
Estimated size of project	Estimated effort < medium effort (33)	0.25	0.10
	Estimated effort >= medium effort (35)	0.31	0.06

TABLE 8
The Three Most Important Reasons Dependent on Data Collection and Analysis Method

Interviews	Experience Reports	Statistical analysis of MRE
No systematic feedback to enable learning	Unexpected events and overlooked tasks	Project estimated by a software developer (as opposed to a project manager)
Poor project planning and management	Change requests from clients or "functionality creep"	Project estimated by a person not participating in the project
Poor requirement specification	Simpler task or more skilled developer than expected (reason for effort under-run)	Client prioritizes time-to-delivery, not cost or quality

5 Conclusion

In the introduction, we stated the *research questions*:

- Are there differences in the reported types of reasons for estimation error which are dependent on the organizational role of the respondent, e.g., whether the respondent is a project manager or general manager?
- Are there differences in the reported types of reasons for estimation error which are dependent on the information collection method, e.g., whether interviews or project experience reports are applied?
- Are there differences in the reported types of reasons for estimation error which are dependent on the analysis method, i.e., whether applying a qualitative analysis of project experience reports or a quantitative (statistical regression) analysis of the same project data?

Our study suggests that the answers to all of these questions are *yes*. We have identified clear patterns with respect to types of reason for estimation errors, which are dependent on respondents' role, data collection approach, and approach to data analyses. This is the main contribution from this paper. We did not identify clearly contradictory reasons for estimation errors when applying different information sources, data collection, and analysis approaches on the estimation error data collected within one company. Instead, we found that the different information sources, data collection approaches, and techniques supported and supplemented each other.

Potentially useful implications from our analysis include the following:

- Identification of indirect reasons (enabling double-loop learning) was much more frequent in general interviews (about 85 percent of the reasons) than in project-specific estimation experience reports (about 45 percent of the reasons), i.e., to get comprehensive reasoning models, a company may need interviews with senior personnel with a general focus on reasons for estimation error and not only project-specific estimation experience reports and questionnaires.
- The identified reasons for estimation *inaccuracy* were described as factors not controlled by the respondent, while reasons for estimation *accuracy* were described as factors within the control of the respondents or related to the respondents' skill or experience. For example, reasons for estimation error provided by the project manager/estimator led to an emphasis on client-related issues, while the interview with the managers of the project managers focused on the need to improve the project managers' skills and processes. This does not mean that the underlying factors, as perceived by the respondents, are very different. Factors described as outside control may in several cases be reformulated as factors inside control, e.g., the same underlying factor may be described as "change requests" (outside project manager control), "poor change request control" (inside control of project managers), or "lack of standards for project management" (inside control of general managers). Our analysis consequently demonstrates the difference in formulation and perspective, not so much that information from different roles leads to identification of different underlying estimation error reasons.
- The use of statistical analyses improved the interpretation and validity of subjective project experience reports and experience-based reasons stated in interviews. This was somewhat unexpected, in light of the low explanatory power and the associative nature of the regression models (regression models are based on covariation and not necessarily on cause-effect relationships).

We recommend that future work on validating our findings should continue to study the impact of whom you ask, how you collect information, and how you analyze the data in different organizations and different project contexts. As part of the validation work, we suggest that controlled experiments should be used. Through controlled experiments, we may be able to better isolate the phenomenon under study. We provide an example of how such controlled experiments on impact of information collection format can be conducted in [21].

We intend to conduct experiments and further observations with the aim of establishing guidelines and frameworks to assist in the development of better and more comprehensive models of factors affecting estimation error. This, in turn, should lead to improved estimation processes and better estimation accuracy. One promising framework, developed for the analysis of errors made in the practice of medical science is described in [22]. That approach is based on predefined levels of reasons, where level 1 reasons describe factors that directly influence the behavior of the individual practitioners, level 2 reasons affect the team-based performance, level 3 reasons relate to the management or organizational level, and level 4 reasons are on a governmental or national level. Other candidate frameworks are Post Mortem Analyses, Ishikawa (fishbone)-diagrams, and Root Cause Analysis.

Acknowledgments

The authors would like to thank Stein Grimstad at Simula Research Laboratory who provided valuable comments and suggestions to this paper.

References

[1] K. Moløkken and M. Jørgensen, "A Review of Surveys on Software Effort Estimation," *Proc. IEEE Int'l Symp. Empirical Software Eng. (ISESE 2003)*, pp. 223-230, 2003.
[2] D. Phan, D. Vogel, and J. Nunamaker, "The Search for Perfect Project Management," *Computerworld*, pp. 95-100, 1988.
[3] M. van Genuchten, "Why is Software Late? An Empirical Study of Reasons for Delay in Software Development," *IEEE Trans. Software Eng.*, vol. 17, no. 6, pp. 582-590, June 1991.
[4] A.L. Lederer and J. Prasad, "Causes of Inaccurate Software Development Cost Estimates," *J. Systems and Software*, vol. 31, no. 2, pp. 125-134, 1995.
[5] G.H. Subramanian and S. Breslawski, "An Empirical Analysis of Software Effort Estimate Alterations," *J. Systems and Software*, vol. 31, no. 2, pp. 135-141, 1995.

[6] A.L. Lederer and J. Prasad, "Information Systems Software Cost Estimating: A Current Assessment," *J. Information Technology,* vol. 8, no. 1, pp. 22-33, 1993.
[7] A. Gray, S. MacDonnell, and M. Shepperd, "Factors Systematically Associated with Errors in Subjective Estimates of Software Development Effort: The Stability of Expert Judgment," *Proc. Sixth Int'l Software Metrics Symp.,* pp. 216-227, 1999.
[8] H.J. Thambain and D.L. Wilemon, "Criteria for Controlling Projects According to Plan," *Project Management J.,* pp. 75-81, June 1986.
[9] D.W.M. Chan and M.M. Kumaraswamy, "A Comparative Study of Causes of Time Overruns in Hong Kong Construction Projects," *Int'l J. Project Management,* vol. 15, no. 1, pp. 55-63, 1997.
[10] H.-T. Tan and M.G. Lipe, "Outcome Effects: The Impact of Decision Process and Outcome Controllability," *J. Behavioral Decision Making,* vol. 10, no. 4, pp. 315-325, 1997.
[11] J. Pearl, *Causality: Models, Reasoning, and Inference.* Cambridge Univ. Press, 2000.
[12] P. Brickman, K. Ryan, and C.B. Wortman, "Causal Chains: Attribution of Responsibility as a Function of Immediate and Prior Causes," *J. Personality and Social Psychology,* vol. 32, pp. 1060-1067, 1975.
[13] K.R. Hammond, *Human Judgement and Social Policy: Irreducible Uncertainty, Inevitable Error, Unavoidable Injustice.* New York: Oxford Univ. Press, 1996.
[14] B. Brehmer, "In One Word: Not from Experience," *Acta Psychologica,* vol. 45, pp. 223-241, 1980.
[15] M. Jørgensen, "A Review of Studies on Expert Estimation of Software Development Effort," *J. Systems and Software,* vol. 70, nos. 1-2, pp. 37-60, 2004.
[16] B. Kitchenham, S.L. Pfleeger, B. McColl, and S. Eagan, "An Empirical Study of Maintenance and Development Estimation Accuracy," *J. Systems and Software,* vol. 64, no. 1, pp. 57-77, 2002.
[17] B. Flyvebjerg, B.S. Holm, and S. Buhl, "Underestimating Costs in Public Works Projects," *J. Am. Planning Assoc.,* vol. 68, no. 3, pp. 279-295, 2002.
[18] A.L. Lederer and J. Prasad, "The Validation of a Political Model of Information Systems Development Cost Estimating," *ACM Trans. Computer Personnel,* vol. 13, no. 2, pp. 47-57, 1991.
[19] C. Argyris, *On Organizational Learning.* Cambridge, Mass.: Blackwell, 1993.
[20] M. Jørgensen, "Regression Models of Software Development Effort Estimation Accuracy and Bias," *Empirical Software Eng.,* vol. 9, no. 4, 2004.
[21] M. Jørgensen, "Realism in Assessment of Effort Estimation Uncertainty: It Matters How You Ask," *IEEE Trans. Software Eng.,* vol. 30, no. 4, pp. 209-217, Apr. 2004.
[22] C. Johnson, "The Causes of Human Error in Medicine," *Cognition, Technology & Work,* vol. 4, pp. 65-70, 2002.

Magne Jørgensen received the Diplom Ingeneur degree in wirtschaftswissenschaften from the University of Karlsruhe, Germany, in 1988 and the Dr. Scient. degree in informatics from the University of Oslo, Norway, in 1994. He has about 10 years industry experience as a software developer, project leader, and manager. He is now a professor in software engineering at the University of Oslo and member of the software engineering research group of Simula Research Laboratory in Oslo, Norway. He has supported software project estimation improvement work in several software companies.

Kjetil Moløkken-Østvold is a PhD student in the Software Engineering Research Group at Simula Research Laboratory in Norway. He received the MSc degree in computer science from the University of Oslo, Norway, in 2002. His main research interests are software effort estimation, management of public software projects, software procurement, and software process improvement.