M. C. Ohlsson, C. Wohlin and B. Regnell, "A Project Effort Estimation Study",
Information and Software Technology, Vol. 40, No. 14, pp. 831-839, 1998.

ELSEVIER

Information and Software Technology 40 (1998) 831–839

INFORMATION AND SOFTWARE TECHNOLOGY

A project effort estimation study

M.C. Ohlsson*, C. Wohlin, B. Regnell

Department of Communication Systems, Lund University, Box 118, S-221 00, Lund, Sweden

Abstract

This paper outlines a four step effort estimation study and focuses on the first and second step. The four steps are formulated to successively introduce a more formal effort experience base. The objective of the study is to evaluate the needed formalism to improve effort estimation and to study different approaches to record and reuse experiences from effort planning in software projects. In the first step (including seven projects), the objective is to compare estimation of effort based on a rough figure (indicating approximate size of the projects) with an informal experience base. The objective of the second step is on reuse of experiences from an effort experience base, where the outcomes of seven previous projects were stored. Seven new projects are planned based on the previous experiences.

The plans are, after project completion, compared with the initial plans and with the data from six out of the seven new projects, to plan the seventh. It is clear from the studies that effort estimation is difficult and that the mean estimation error is in the range of 14%–19% independent of the approach used. Further, it is concluded that the best estimates are obtained when the projects use the previous experience and complement this information with their own thoughts and opinions. Finally, it is concluded that data collection is not enough in itself, the data collected must be processed, i.e. interpreted, generalized and synthesized into a reusable form. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Empirical study; Experience base; Effort; Data collection; Measurements; Education

1. Introduction

Effort estimation is a crucial task for the planning and pricing of software development. There exist a number of methods that, based on historical data and problem size, predict development effort, e.g. COCOMO [1] and Putman [2]. Comparisons of different effort estimation models is presented by Kemerer [3] and Kitchenham [4].

This paper presents an emprirical study of the stepwise introduction of an effort experience base, in accordance with the notion of an experience factory [5, 6] or a corporate memory [7]. The main idea is to reuse experience from earlier projects as a foundation for process improvement. The experience factory is a separate organization, which collects and analyses data from the development organization and packages the experience into a reusable experience base. This synthesized and packaged experience, supports the development organization in improving, e.g. effort estimation.

The objective of the presented study is to evaluate if an effort experience base improves estimation accuracy. Here, we present results from the first two steps in a larger, four-step, empirical study. The study is carried out in the context of a stepwise formalization of the effort experience base. The steps differ in the degree to which the experience base is formalized. In addition, each step applies a more formal way of collecting and analysing effort data. The study is carried out in an educational environment, where 7 groups of 14–19 students independently develop telecommunication systems based on the same requirements definition.

In the first step, the input is a 'rough figure' and the output is an 'informal experience base', shown later. The evaluation is done by comparing the predictability based on the input and the predictability that we would have expected if the output would have been available at the planning stage of the project. This is further elaborated later. The four steps can be described as follows:

Step 1. Rough figure → informal experience base (fall 1995)
No experience base was available when the projects started.

Step 2. Informal experience base → template based experience base (fall 1996)
The experience base from the first step is provided, although without any particular training of using it. Templates for reporting data consistently are provided. The objective is to evaluate if the outcome is better than in Step 1 and also to

* Corresponding author. Tel.: + 46 46 222 0367; fax: + 46 46 145 823; e-mail: magnuso@tts.lth.se

PII: S0950-5849(98)00097-4

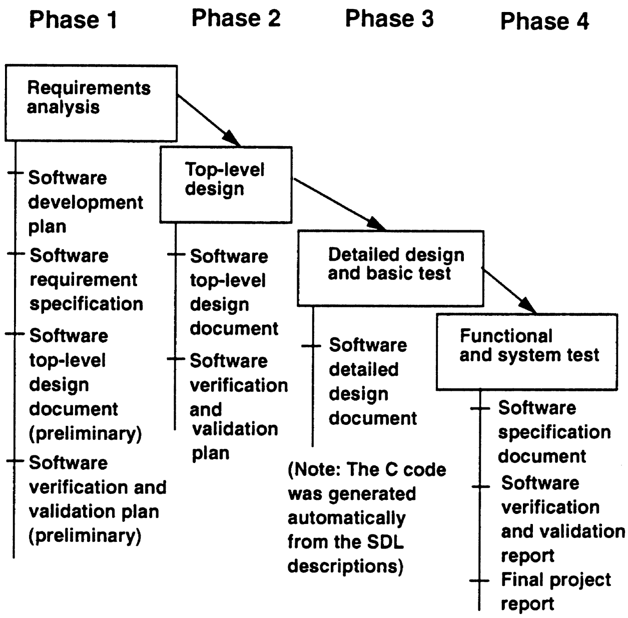

Fig. 1. The waterfall development model.

add more projects to the experience base and perform a similar evaluation as in Step 1. This step is further described in the next section.

Step 3. Template based experience base → formalized experience base (fall 1997)
In the third step, the experience base from the projects in Step 2 is made fully available. The project managers will receive training and information concerning project planning using an experience base. The analysis is not made yet.

Step 4. Formalized experience base → extended formalized experience base (fall 1998)
The final step includes using the experience base and evaluating if the predictability of effort improves as the formalization of the education concerning effort estimation using an experience base increases.

After the four steps have been conducted, the objective is to recommend how to create an effective effort estimation experience base. In particular, the aim is to recommend a suitable level of formalization of the experience base and to make the data in the experience base useful for planning purposes.

The paper is organized as follows. Step 1 and 2 of the study is described in Section 2 and the data analysis is described in Section 3. Conclusions are presented in Section 4.

2. Study outline

2.1. Objectives

We have two objectives of the study. They are to evaluate the usefulness in using historical data for these estimates and to evaluate the needed formalism to improve effort estimation from an managerial viewpoint. From these we can formulate the following questions:

- Does the use of historical data improve the effort estimates?
- Does increased data collection formalism improve the effort estimates?

Also, the intention is to evaluate if a rough and straightforward figure (historical data without any modifications) would provide an estimate as good as the use of the experience data.

2.2. Context

The subjects, i.e. the projects used in the study are chosen from a software development course conducted at the department of Communication Systems, Lund University. The students were unaware of the intention to use the information for the study. The software development course is read by over 100 students yearly, which means that 6–8 similar projects are run in parallel since each project is manned with 14–19 students. Each project develops a number of telephone services [8].

2.2.1. Process model

A somewhat simplified version of the DoD-2167A [9] is used. It is a waterfall model and the objective is to provide a standardized approach, rather than experimenting with a more sophisticated process model. The model, as used in the project, is outlined in Fig. 1. In particular, the baselines in the model are emphasized to provide checkpoints in the process, see also the next subsection. The experiment described here is focused around effort needed and used for the different phases in the process model.

2.2.2. Project

The groups consist ideally of 16 students divided into 8 subgroups. One subgroup acts as project leader, one subgroup is a system group responsible for keeping the system together, four subgroups are development groups, where each group develops one service and finally two subgroups are test groups testing two services each and with a joint responsibility for the system test (see Fig. 2).

The roles of the subgroups are combined with roles played by department personnel. The department personnel play the roles of customer, external quality assurance personnel and, of course, technical experts, who can help the students when they run into problems. The customer reviews the produced material at two times during the project and also performs an acceptance test at the end of the project. It should, however, be noted that executing the

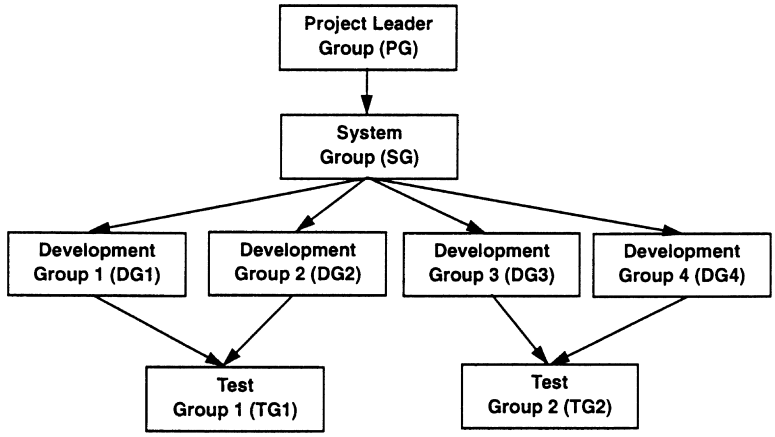

Fig. 2. Organization of the projects

process correctly is viewed equally important as the final product.

2.2.3. Conducting a project

The first two weeks are quite hectic with a number of lectures giving the students an introduction to the development environment and the project. The students are allowed to form groups on their own and the project leader subgroup is appointed. This subgroup is then responsible for the division into the other subgroups. This should be settled after the first week and they should familiarize themselves with their work in the forthcoming weeks.

The requirement analysis documents should be written during the second week, they should be ready for a customer review during the third week of the project. The requirement specification should be a baseline after the review. In the third and fourth weeks of the project, the students are mainly working on the top-level design document. This document is reviewed by the customer in the fourth week, which means that the customer can insure that the project is heading in the right direction.

The latter part of week four and into week six is focused upon the detailed design document and the product should be delivered in the end of the sixth week. A project report should be delivered together with the product. This report should contain information about the work and in particular metrics should be reported concerning time expenditure and faults found during inspections and testing. The delivered product is acceptance tested in the seventh week and the students are given feedback on their performed work.

2.3. Design

2.3.1. Step 1

The first step of the study is based on seven projects conducted by 108 students in 1995. The seven projects were exactly the same, with the exception that two projects were reduced slightly in size to accommodate the two project groups which only consisted of 14 students, while the other five projects had 16 participants. The ability to run seven very similar projects in parallel provides a unique basis for evaluating the experience approach. If the effort experience base does not work when the projects are almost identical, then it is hard to believe that it will work in an industrial environment (where hopefully all projects are unique in some way). It should, however, be noted that the benefits of an effort experience base cannot be fully evaluated until all four steps of the study were carried out.

2.3.2. Step 2

The second step of the study was carried out based on projects executed during the fall of 1996. The number of students attending the course was 123. They were divided into 5 projects with 17 students and 2 projects with 19 students.

The students obtained the informal experience base, based on the projects conducted in 1995. The experience base is considered to be informal, since it is synthesized from data collection based on informal procedures, for example, no data collection templates and guidelines were used. The objective was that the experience base should help

Table 1
Total time for each type of group, EB95 (Experience Base 1995)

Measures	Project group (PG)	System group (SG)	Development group (DG)	Test group (TG)	Meeting percentage overhead (%)
Mean	57.2 h	123 h	67 h	94.9 h	39.6
Standard deviation	16.5 h	28.1 h	20.2 h	29.2 h	14.4

Table 2
Distribution of effort between the different phases in percentage, EB95

	Phase 1 (%)	Phase 2 (%)	Phase 3 (%)	Phase 4 (%)
PG	65.1	5.8	3.7	25.5
SG	35.5	35.2	26.6	2.7
DG	31	27	41.1	0.9
TG	34.2	31.3	0	34.5
Meetings	29.2	26.5	23.7	20.7
Overall	34.4	27.2	22.6	15.9

the students to plan their project. The information provided is given in Table 1 and 2.

In Table 1, the number of man-hours for each type of group is given. It is a mean calculation of how much effort the groups have put into the projects. The meeting percentage is calculated as the percentage of the total project time, i.e. according to how much effort the groups use, there is another 39.6% that should be added to the total effort. This figure is relatively large and the reason for this is that many of the projects did not really know how to classify the time spent on, for example, review meetings.

The students were asked not only to focus on these figures when estimating their projects, but to use this information as a complement to their own view of effort required to perform the project and the different tasks in the project.

Table 2 shows the division of the effort between different phases in the software development process. The phases are primarily related to documents to be produced and not to when the work is carried out in time (see Section 2).

2.4. Threats

All studies are threatened in one way or another. It is always difficult to conduct studies where the results can be interpreted correctly (internal validity) and an even more difficult task is to interpret the generality of the results (external validity). In this particular case the following main threats were identified.

- Instrumentation

  No particular instrumentation concerning data collection was carried out in Step 1. However, we believed that, based our insight into the projects during the execution, that we are able to make rather good interpretations.

- Student setting

  Experiments in a student setting can always be questioned concerning validity in an industrial environment. In this case, this is not regarded as particular critical as one objective of the course is to model an industrial environment. In particular it should be noted that the study is based on comparison of different methods for effort estimation and the evaluation should provide similar results independent of the environment (university or industry).

- Reliability of the data (the students may not be reporting properly)

  The progress reporting weekly is believed to make these risks rather small.

- Division into project groups and lack of actual instrumentation

  The division into groups is made by the students and it may lead to some groups knowing each other better than others.

2.5. Operation

All the data used in the calculations come from the projects final reports. These reports contain information about how much time the projects have spent on different activities and the fault density for different documents, based on faults that were found during inspections and tests. There are no precise specifications for how the data should be reported. Therefore, the granularity of the reported data differs from project to project.

In 1995 no special forms were provided to help the projects collect data. The projects had to take care of this themselves and create their own procedures and forms.

The data in the reports from 1996 are a fusion of data from other reports with finer granularity. Daily the projects members noted how much time they had spent on different activities. This information was gathered in an individual time report and this data formed the base for the weekly report. The weekly report was filled in at the end of every week. The weekly report should contain a summation of the worked hours for the current week. At the end of the projects the weekly reports were gathered and summarized in a final report, from which we, as mentioned previously, collected our data.

3. Data analysis

The intention of the analysis is to see if it is possible to make better time estimations, based on experiences from earlier projects and to validate if the error in planning has decreased from Step 1 to Step 2 by collecting the data in a more formalized way.

3.1. Analysis strategy

The objective is to evaluate if the estimation using the experience base approach is better than the initial estimates made by the students for the projects. The long term objective is to carry out the other steps as planned in Section 1 and evaluate if the estimates become better as the effort experience base is more and more formalized.

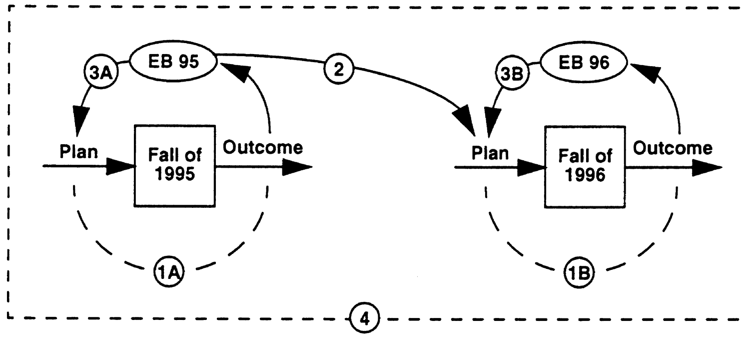

Fig. 3. Analysis strategy overview.

The analysis could be summarized in four steps (see Fig. 3):

1. Plan vs actual: a comparison of the projects' estimated time vs the actual (1995 and 1996, 1A and 1B).
2. Straightforward experience base: a comparison of the actual effort vs a summation of the figures from the experience base in Step 1 (1996). The summation figure is a calculation made of the figures in the experience base without any special treatments applied. This figure is used as a plan for the project.
3. Experience base in retrospect: we have used six out of the seven projects to create an experience base in retrospect. This is performed for all seven projects and the figures from these calculations are compared to the actual figures (1995 and 1996, 3A and 3B).
4. Improvements of the estimations: a comparison of the relative errors for the estimates to see if they have decreased from Step 1 (1995) to Step 2 (1996).

The analysis can easily be done for mean values, but it is also possible to determine intervals, which could be useful in project planning in order to determine an effort interval assuming a normal distribution and using the *t*-distribution. We have here, however, chosen to focus on an ANOVA test as, the main objective is to evaluate the predictability of different estimating methods from the three first analysis steps mentioned previously. In particular, in Step 3 we would like to compare the prediction made with the rough figure (the plan of the students) with the one we could have obtained if the other six projects would have been run before the project we are currently assuming to plan (a hypothetical plan based on the informal experience base). For the last analysis step we will apply a *t*-test to verify if there exists any significant difference.

3.2. Analysis

3.2.1. Plan versus actual (1A and 1B)

In 1995 the projects received a figure, for planning, telling that the projects are expected to take approximately 1000 man-hours. But at the start of the projects in 1996, the participants received the figures from Tables 1 and 2 (based on the outcome from the projects in 1995) and used them to plan the projects. We have compared the estimated time with the actual outcome to find out how accurate their estimates were. The difference is calculated as *(Plan-Outcome)/Outcome* and the results can be found in Table 3 and 4.

The outcomes were acceptable even if some of the projects had large deviations. Some of the project groups planned their projects unexpectedly well.

3.3. Straightforward experience base (2)

If the projects had estimated solely with the figures from the experience base (Tables 1 and 2) without any modifications the estimates and the difference in the estimates would have been as presented in Table 5. The difference is calculated as *(EB-Outcome)/Outcome*.

The reason for doing this is to study if it would have been better if the estimates were made straightforward and the project groups did not spend time on planning the projects in detail. Also, we can see how useful the raw data from the experience base is.

The calculation for Project 1's EB value of 891 h can be found in Table 6. Project 1 consisted of four development groups and two test groups. Therefore, the values from Table 1 have to be multiplied with four and two, respectively. Finally the overhead time for meetings is added (cf.

Table 3
Outcome vs plan 1995

	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7
Plan 95	1182	886	755	690	965	1001	1004
Outcome 95	963	599	745	810	1030	1000	894
Difference	22.7%	48%	1.3%	−14.8%	− 6.3%	0.1%	12.4%

Table 4
Outcome vs plan 1996

	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7
Plan 96	1346	1415	1172	1010	1183	1200	1096
Outcome 96	977	1385	958	1021	1265	1630	951
Difference	38%	2.1%	22%	−1%	− 6.5%	−26%	15%

Table 5
Outcome vs straightforward EB calculation 1996

	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7
Outcome 96	977	1385	958	1021	1265	1630	951
Straightforward EB 96	891	984	891	891	891	984	891
Difference	−8.9%	−29%	−7.1%	−13%	−30%	−40%	−6.4%

percentage figure in Table 1). This calculation was made for the other six projects too.

The result from this analysis is that some of the projects' plans would have had quite large deviations in comparison with the outcome. The three projects with the largest deviations would have been Projects 2, 5 and 6. Projects 2 and 6 are the only two that had five development groups instead of four and Project 5 had made a very detailed software development plan. From Projects 2 and 6, we can see that adding an extra development group also adds management time.

3.3.1. Experience bases in retrospect (3A and 3B)

The objective of this analysis is to evaluate if the estimates of the projects would have become better than the initial estimates if they had used the outcome from the projects, run in parallel, as an experience base. This means that we have experiences from projects which were run under the same circumstances as the one we should plan. Also, this analysis forms a base for the evaluation of the four step study, i.e. to evaluate if the estimates and plans become better as the experience base becomes more formalized.

For all seven projects in 1995 and in 1996, we have created an experience base, based on the other six projects from the same year. This is done in a 'circular' way. The creation of the experience bases is based on four parameters.

- Mean value and standard deviation for each group type. We have focused on mean values in our analysis because our main concern is to evaluate predictability, even though it is possible to determine intervals and use them in planning to determine intervals of effort by assuming a normal distribution and using the *t*-distribution.
- Division of effort in percentage for each group type between the different phases.
- The time is connected to the phases, which in turn are connected to different documents and inspections.
- The meeting time is added as a percentage figure because of the varying amount of people in the projects.

As mentioned earlier, the final reports from the projects differ a bit. Therefore, some corrective actions were taken. One example is Project 5 in 1996 which has reported a summation of the time worked in the phases, including time that is not interesting for the study. For example individual study times and lectures.

To estimate one project we use the other six to create an effort experience base in retrospect. This is done for all the seven projects and the relative differences between the outcome of the projects and the retrospect experience base are presented in Table 7 and 8.

In Table 7 from 1995 we can see that it is only two out of the seven projects that have a difference smaller then 10% and three of them had very large deviations. Project 2 and Project 5 consisted of three development groups and they have extremely different outcomes. Even though they have produced the same software the figures vary very much.

Four out of the seven projects in Table 8 have a difference smaller than 10% and the deviations are not as large as the ones from the straightforward experience base. The largest differences are still related to Project 2 and 6, which consisted of five development groups.

3.4. Improvements of the estimations (4)

To be able to see the outcome of the study, we have to compare the figures from Step 1 and the new figures from

Table 6
Project plan for Project 1 1996

Project 1	Total
PG	57.2
SG	123
DG (4 subgroups)	$67 \times 4 = 268$
TG (2 subgroups)	$2 \times 94.7 = 189.8$
Meetings	$(57.2 + 123 + 268 + 189.8) \times 0.396 = 252.7$
Total	891

Table 7
Outcome vs Retrospect EB, 1995

	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7
Outcome 95	963	599	745	810	1030	1000	894
Retrospect EB 95	879	833	919	904	759	865	890
Difference	−8.8%	39.2%	23.4%	11.7%	−26.3%	−13.5%	−0.4%

Table 8
Outcome vs Retrospect EB, 1996

	Project 1	Project 2	Project 3	Project 4	Project 5	Project 6	Project 7
Outcome 96	977	1385	958	1021	1265	1630	951
Retrospect EB 96	1027	1079	1025	1024	987	1054	1035
Difference	4.6%	−22.5%	6.4%	0.3%	−22%	−35.6%	8.3%

Table 9
Comparison of improvements

Step 1 and 2	Project 1 (%)	Project 2 (%)	Project 3 (%)	Project 4 (%)	Project 5 (%)	Project 6 (%)	Project 7 (%)	Mean (%)
Rough figure 95	22.7	48	1.3	−14.8	−6.3	0.1	12.4	15.1
Retrospect EB 95	−8.8	39.2	23.4	11.7	−26.3	−13.5	−0.4	17.6
Informal EB 96	38	2.1	22	−1	−6.5	−26	15	15.8
Straightforward EB 96	−8.9	−29	−7.1	−13	−30	−40	−6.4	19.2
Retrospect EB 96	4.6	−22.5	6.4	0.3	−22	−35.6	8.3	14.2

Step 2. Also, it is necessary to compare if the idea with the experience bases work or if it is enough with a rough effort figure when planning projects. The figures for this comparison can be found in Table 9. The figures in the table come from the projects run in 1995 (forming Step 1) and from the projects run in 1996 (forming Step 2). When we calculated the mean relative difference, we did not take into account if it was an over or underestimate, i.e. we used the absolute value.

The different estimating approaches from 1996 show that the informal experience base is best in four out of seven cases, the straightforward experience base in two and the retrospect experience base in one case, even though the figures for Project 4 only differs 1.3% points between the informal experience base and the retrospect experience base. From the study in 1995, we can see that the rough figure is best in four cases out of seven.

The results from the statistical analysis did not provide us with significant results. This disabled us to confirm any of our questions with significance, according to the objectives of this study. Though, the studies from 1995 and 1996 indicate that the best estimates are actually the ones the students made themselves, both in the study with a rough figure in 1995 and with an informal EB in 1996. This shows the importance of adding personal opinions to any data you have in an experience base. Consequently, this implies that we are not able to produce better estimates using the available data than the students do when they plan their projects from the rough figure provided to them.

For projects with four development groups and two test groups, the experience base seems to be useful, but when planning projects with three or five development groups, much more concern has to be focused on the overhead when the projects become smaller or larger, respectively. Project 2 (1996) managed to do this and their relative error was just 2.1% while the straightforward experience base error was 29% and the retrospect experience base error was 22.3%. Still, none of these figures from the retrospect experience base are fully acceptable, especially from a management perspective.

Instead, if we compare the mean value for the difference, we can see that the span between them is small and it is just a minor improvement for the 1996 retrospect experience base in comparison with the 1995 retrospect experience base. The mean values for the relative errors differ between 14.2% and 19.2% and this is not large enough to pinpoint one of the approaches to be superior.

Another important aspect to analyse, is how the distribution of the time between the phases and groups have changed. The total time for each type of group from the projects in 1996 is presented in Table 10. It is worth noting that the system groups 1996 consisted of three persons and in 1995 they consisted of two persons, though, the workload was still the same (compare the figures in Tables 10 and 1).

The effort distribution between the phases for 1996 years projects is presented in Table 11.

Tables 10 and 11 should be compared with Tables 1 and 2 from 1995 and, hence, Tables 10 and 11 are the effort

Table 10
Total time for each type of group, EB96

Measures	Project group (PG)	System group (SG)	Development group (DG)	Test group (TG)	Meeting percentage overhead (%)
Mean	97.7	235	98	119.6	14.7
Standard deviation	36.3	75.1	17.6	35.1	3.3

experience base created from the seven projects run in 1996. The use of templates gave us the ability to trace the effort distribution, between the phases and the groups, better. For example, the meeting percentage has decreased, i.e. the students were able to distribute the man-hours among the phases to a larger extent.

If we add the meeting overhead percentage to the different types of groups in Tables 1 and 10 (e.g. PG95 $57.2 \times 1.396 = 79.9$ and PG96 $97.7 \times 1.147 = 112.1$) the total effort per group has increased. Finally, the distribution between the different phases has not changed significantly.

4. Conclusions and future work

This paper has outlined a four step study with a successively more and more formalized way of estimating project effort. The objective of the study is to evaluate the needed formalism to succeed with a project effort experience base. The background to the study is that many companies collect data today (with the hope to be able to reuse it later on), but few of them collect consistent data throughout the whole company. In particular, there is generally a lack of consistent data to be able to generalize and make the data reusable by software projects to come. Thus, the motivation for the study is simple, there is a need to illustrate the formalism needed to succeed with creating a useful experience base. Data collection as such is certainly not enough, which was shown by the first step of the study. Reuse of project experience needs more than just data collection in general to be useful, otherwise, general figures of experience (project leader experience) are as useful as the support obtainable from a data repository.

The study is conducted in quite a unique environment in the sense that a number of projects with the same problem description are performed. This creates an interesting environment to evaluate an experience base. If the experience base (with a certain level of formalism) does not work in this environment, it is not likely that it will work in an industrial environment. Further, this study is believed to be the first of its kind.

We have shown that a rough figure is just as good as a figure from the experience base but the idea of an experience base should not yet be ruled out. Instead, we have to improve the collection of data and the granularity of the software development plans. If we do this the figures will probably be more accurate and it will help us interpret the data to be able to analyse it. An approach could be to use an experience base combined with expert judgement.

Therefore, in Step 3 the project managers will be trained in using and maintaining the experience base and they will be able to see how the data is used and this will hopefully lead to better planning and reporting. Also, we will provide them with new templates, for reporting the desired data in the final reports.

The results so far give some initial indications regarding the effort experience base. From the first step, it was concluded that data collection in itself is not enough to succeed with an experience base. Now, in the second step a small improvement was achieved, although the improvements are not statistically significant. The relative errors in the planning of the projects have decreased and the control of the effort has increased. This could be interpreted as supporting an effort experience base, but so far we have not proven that it is useful.

Another important issue is, of course, to challenge others to perform similar studies in order to improve the understanding of reuse of data for project planning. Experience in terms of studies concerning experience bases is vital for companies in order to identify a cost-effective way of implementing an experience factory in a specific environment. Thus, we are looking forward to other studies in the area, and to case studies addressing these issues.

Acknowledgements

We would like to thank all the students in the course who put in a great deal of effort and provided valuable feedback to us both in terms of data and comments on the course. Also, we would like to thank Martin Host, the Department of Communication Systems, for valuable comments on previous versions of this paper. This work was partially funded by the Swedish National Board for Industrial and Technical Development (NUTEK), P 10495-1.

Table 11
Distribution of effort between the different phases in percentage, EB96

	Phase 1 (%)	Phase 2 (%)	Phase 3 (%)	Phase 4 (%)
PG	50.9	9.2	10.8	29
SG	25.3	41.2	30.1	3.4
DG	19.1	32.4	43.4	5
TG	53	16.8	2.7	27.5
Meetings	25.2	27.1	20.6	27.2
Overall	30.9	27.8	26.8	14.5

References

[1] B.W.Boehm, Software Engineering Economics, Prentice Hall, Englewood Cliffs, NJ, 1981.

[2] L.H. Putman, A general empirical solution to the macro software sizing and estimating problem, IEEE Trans. Software Engng SE-4 (1978) 345–361.

[3] C. Kemerer, An empirical validation of software cost estimation models, Commun. ACM 30 (1987) 416–429.

[4] B. Kitchenham, Software development cost models, in: P. Rook (Ed.), Software Reliability Handbook, Elsevier, Amsterdam, 1990, pp. 487–517.

[5] V. Basili, G. Caldiera, H.D. Rombach, Experience factory, in: J.J. Marciniak (Ed.), Encyclopedia of Software Engineering, vol. 1, Wiley, New York, 1994, pp. 469–476.

[6] F. McGarry, Porting experience factory concepts to new environments, in: Proceedings of the Twentieth Annual Software Engineering Workshop, NASA, Greenbelt, Maryland, USA, 1995, pp. 71–92.

[7] B. Lennselius, C. Wohlin, C. Vrana, Software metrics: fault content estimations and software process control, Microprocessors Microsystems 11 (7) (1987) 365–375.

[8] C. Wofflin, The challenge of large scale software development in an educational environment, in: Proceedings Conference on Software Engineering Education and Training, Virginia Beach, Virginia, USA, 1997, pp. 40–52.

[9] DoD-STD-2167A, Defence System Software Development, 1988.