Developing A Usability Capability Assessment Approach through Experiments in Industrial Settings

Timo Jokela*, Marko Nieminen†, Netta Iivari*, Katriina Nevakivi‡ & Mikko Rajanen*

* *PO Box 3000, 90014 University of Oulu, Finland*
Email: *{timo.jokela,netta.iivari,mikko.rajanen}@oulu.fi*

† *Helsinki University of Technology, Otakaari 4, 02150 Espoo, Finland*
Email: *marko.nieminen@hut.fi*

‡ *Buscom, Elektroniikkatie 4, 90570 Oulu, Finland*
Email: *katriina.nevakivi@buscom.fi*

Usability capability assessments are carried out to analyse the capability of a development organisation in performing user-centred design (UCD). We carried out four experimental usability capability assessments to learn how to perform assessments effectively in industrial settings. Our starting point was traditional software process assessment based on ISO 15504 ('SPICE'). The recent ISO/TR 18529 was used as the process reference model of UCD. Our experiments showed that the focus of ISO 15504 process assessments — management of activities — did not exactly meet the needs of assessments in our context. These experiences led us to a modified assessment approach where the focus is in performance of UCD. Its main characteristics are:

1. **a refined UCD process model;**
2. **a three-dimensional capability scale; and**
3. **implementation of an assessment as a workshop rather than a series of interviews.**

Keywords: UCD, usability capability, usability capability assessment, usability maturity models.

1 Introduction

The challenge to improve the position of UCD (UCD) in development organisations has been recognised in many presentations and panels in conferences and seminars. For example, there have been papers (Rosenbaum et al., 2000), tutorials (Bloomer & Wolf, 1999), panels (Rosenbaum, 1999) and interviews (Anderson, 2000) at CHI conferences. A European TRUMP project has also addressed this topic.

A typical approach to start organisational improvement efforts in any domain is to carry out *current state analysis*. Through current state analysis, one can identify the strengths and weaknesses of an organisation, and thus get a good basis for planning and implementing improvement actions. For example, in software engineering, current state analysis is a widely used practice in the form of *process assessment*. Recognised process assessment approaches are *CMM* (Paulk et al., 1995), *Bootstrap* (Kuvaja et al., 1994), and *ISO 15504* (ISO, 1998a). When we examine the recognised *process improvement models* of software engineering, for example IDEAL of Software Engineering Institute (McFeeley, 1996) and ISO/TR 15504-7 (ISO, 1998b), they essentially include process assessment as a step in an improvement process.

In UCD, similar activity called *usability capability assessment* (UCA) seems to gain popularity. According to a study of Rosenbaum et al. (2000), 16 organisations out of 134 (12%) reported using 'organisational audits' as a means for enhancing 'strategic usability'. *Our research problem is to learn how to perform usability capability assessments effectively in industrial settings.*

In this research, our hypothesis was the traditional software process assessment, as defined in ISO 15504[1], using the recent *ISO/TR 18529* (ISO, 2000) as the process reference model of UCD. A pre-version of the model is the *UMM Processes* (Earthy, 1999). It was originally developed in the European *INUSE* research project and further elaborated during another European research project, *TRUMP*. In the beginning of our research, ISO/TR 18529 was not yet approved, and our reference was the UMM Processes model. — In this paper, however, we use consistently the term ISO/TR 18529.

ISO/TR 18529 is a process model developed specifically for process assessment. The format of its process definitions complies with the requirements of ISO 15504. Altogether, ISO/TR 18529 identifies seven UCD processes. Five of them are derived from the standard ISO 13407 (ISO, 1999). The processes are further divided into base practices. In an assessment, the capability of a process is typically determined through performance of *base practices*. The result of an assessment is a *capability profile*: each UCD process is given a capability rating. There are six levels of *capability*, from 0 (lowest) to 5 (highest). — The processes of ISO 13407 are illustrated in Figure 1. ISO/TR 18529 is discussed more detailed in Bevan & Earthy (2001).

[1] Also known as 'SPICE'

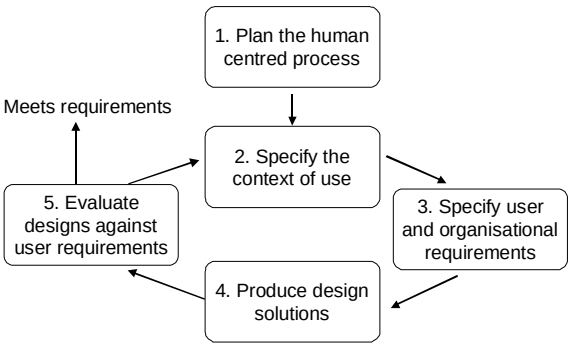

Figure 1: Activities of UCD as defined in ISO 13407.

2 Overview of the Assessments

We carried out four experimental usability capability assessments in three industrial companies in Finland. We started at *Buscom* with a traditional process assessment by using the ISO/TR 18529 process model as a reference in May 2000. Thereafter we performed an assessment at *Teamware* in June 2000. The third assessment (NET 1) was carried out at *Nokia Networks (NET) IMN* user interface team in October-November 2000. The biggest methodological step took place during this assessment. We developed revised process and capability models based on the experiences from the two previous assessments. The fourth assessment (NET 2) was at Nokia Networks in a customer documentation team in December 2000. This time we implemented the assessment as a workshop rather than as a series of interviews. The flow of the assessments is illustrated in Figure 2.

We describe each of the four assessments one by one, pointing out the main characteristics of the assessments and lessons learned. We describe the basics of the evolved models as part of the discussion of the NET 1 case.

3 Assessment at Buscom in May 2000

Buscom develops fare collection systems for transportation systems — especially for bus companies. The company has two features that make an assessment a special case. First, the company is small — about 60 employees. Another specific feature is that at the time of assessment, the company had only limited background in UCD. Still the company found it sensible to start development of UCD processes with an assessment.

3.1 Implementation of the Assessment

There were altogether six members in our assessment team. Two members of the team were trained process assessors. One of the team was the usability person of the company.

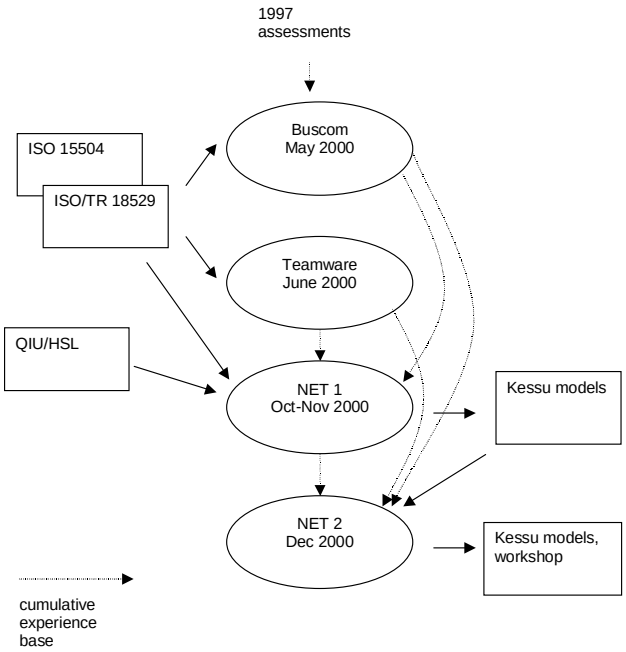

Figure 2: Summary of the assessments.

Based on our previous experience from assessments in 1997, reported in (Kuutti et al., 1998), we decided to assess to the level 1 only. In other words, we focused on assessing the substance of UCD — not on how the activities are managed. We interviewed the representatives from a number of development projects during two weeks' time, having altogether 10 interview sessions. There were one, two or three interviewees in each interview session. The total amount of interviewed persons was 20. On many days, we had two interviews per day.

We used base practices of the ISO/TR 18529 model as our reference in the interviews. We put effort in understanding what is meant with each base practice and in understanding to which extent the base practice was performed in the development process of the organisation. To do this, we assigned a responsible person for each process in our assessment team.

Rating the capability of processes, however, was not as successful as we had planned. During the assessment, we decided to give up giving ratings on the basis of base practices. The reasons for this were partly the tight schedule, partly the difficulties in the interpretation of base practices. The assessment team felt that it was difficult to give valid capability ratings. Our decision was the rate the processes informally, based on professional judgement rather than on the based practice driven capability-rating algorithm. We reported the results emphasising the qualitative findings.

3.2 Feedback from the Assessment

We gathered feedback from the assessment by using questionnaires delivered to the audience in the results presentation session. The staff reported that they generally found results of the assessments useful. A clear result was that those who were interviewed learned more about UCD than those that attended the result presentation session only. The core processes of ISO 13407 were perceived understandable. The results reporting session that lasted one hour, however, was perceived somewhat 'boring'.

In addition to gathering feedback with questionnaires, we later conducted some in-depth interviews of some key persons. The interviews revealed some interesting points that were not covered by the questionnaires. For example, we had presented a definition of processes and the related base practices in the form of ISO/TR 18529 model. The usability person reported afterwards that the interviewees had found the definitions of the processes and base practices difficult to understand. Some of the staff had a perception that the assessment was "academic stuff driven by the interests of the university".

The usability person of the company said that the assessment was personally a very positive learning experience. She learnt a lot not only about assessments but also got a more thorough understanding of the practices and principles of UCD and of the current ways of working in different units of the organisation. The assessment team found generally the UCD processes of the model sensible. The main problem was, as said, in the interpretation of the base practices.

3.3 Lessons Learnt

Based on these experiences it seems that assessment is not only about identifying strengths and weakness but also about communicating about usability and UCD to the organisation. We did not find as a very good result that many members of the staff had perceived 'not understandable' the things that we had presented. We conclude that improvement in the assessment process should be done in the following areas:

- From the assessors point of view, the main lesson learnt was that more precise and unambiguous interpretations of the base practices of the processes are required.
- The assessment should be carried out in slower pace. Too frequent interviews did not allow time to 'stop and think'. Two interviews a day may be appropriate for routine assessment, but not for research.
- The assessment process should be planned better and scope of the assessment should be only a very limited number of projects.
- Communicating the basics of UCD and the results of an assessment is a challenge. Results should be presented both qualitatively and quantitatively in a concise form.

The positive lessons learnt were:

- Limiting the assessment to level 1 of capability was the right choice: it made sense to examine only the essential performance of UCD. Assessment of management issues (levels 2 of capability and above) had not been meaningful in this case.

- The basic concepts of ISO 13407 (processes, definitions of usability etc.) are useful. They give a good basis for assessments.

- Interviews were experienced to be effective learning processes by the interviewees and the usability person.

4 Assessment at Teamware in June 2000

Teamware Group is an international software development company with several years' experience in UCD. Its main business is Internet based solutions for communities. A UI design team within the company, Ergo Team, has operated for years, and has influenced remarkably in the improvement of the user-centred development practices in the company.

This assessment followed an assessment that was conducted at Teamware in November 1997, reported in Kuutti et al. (1998). Even if we found some problems with the ISO/TR 18529 model in the assessment at Buscom, the approach was not changed for this assessment. One reason for this was that Teamware wanted to have comparative results with the assessment that was carried out earlier. Another reason was that the lead assessor was from a different organisation.

4.1 Implementation of the Assessment

The assessment team consisted of seven persons. Additionally, one representative of the assessed organisation was present in all assessment sessions. Most members of the assessment team had participated software process assessment training according to Bootstrap method (Kuvaja et al., 1994) after the assessment at Buscom.

All the processes of ISO/TR 18529 (HCD.1-7) were on the focus of assessment. The assessment method was traditional: the capability of the processes was determined using the base practices. The goal was to assess the processes up to level 3 of process capability if applicable — as was done in 1997. The focus of the assessment was in the early phases of development although all processes were assessed. The customer defined the focus.

The assessment lasted one week. Eight interview sessions (nine persons) took place during the week. The results were reported to the representatives of the organisation on the last day.

4.2 Feedback from the Assessment

We delivered questionnaires to the audience in the results presentation session. The questionnaires revealed that most interviewees felt that the interviews handled meaningful issues. Some pointed out that due to the insufficient information provided by the opening briefing they could not prepare themselves well enough. Most of the interviewees reported on gaining new ideas concerning their work. However, managers felt that the interviews did not handle very meaningful issues.

In results reporting session, the respondents considered all the UCD processes to be very important — if not for them, then for the company. Otherwise, they criticised the assessment results. They felt that many important areas related to the UCD were not discussed at all in the interviews. They felt that model had limited discussions sometimes to even irrelevant topics. Consequently, some felt that the results did not describe reality very well. The respondents also criticised that we did not explain well enough the terminology used or the maturity scales presented. We presented the results qualitatively, but the audience wished for the qualitative results, too.

The assessment team experienced the assessment week rather frustrating. The biggest problem was — as was at Buscom — in the interpretation of base practices. The interpretations caused even more disputes within the assessment team than at Buscom — now there were members from two organisations in the assessment team. Especially, there were disagreements whether a process truly reaches level 1, and wheter it makes sense to examine upper levels of capability. Some members of the assessment team experienced a problem in the validity of the interview style: interviewing the processes through base practices one by one. Many members of the assessment team felt that they did not get a good picture of essential practices of the company. These problems also led to difficulties in rating the capability of the processes. The capability scores were given but the interpretation of the findings remained contradictory.

We show some examples of base practices in Table 1. We had trouble in agreeing on interpretation with practices such as “Analyse the tasks and worksystem” and “Analyse the implications of the context of use”. One the other hand, there were no problems to interpret a base practice such as “Describe the characteristics of the users”.

4.3 Lessons Learnt

We got confirmation to our understanding that an assessment is not only about identifying strengths and weaknesses but also about increasing the awareness and commitment of the personnel towards UCD. Interviews were found to be an effective learning process to the interviewees, at least to the designers. Also the experience of the company from the previous assessment confirms those who were interviewed got committed to UCD.

Most of the lessons learnt are in line with the lessons learnt from the assessment at Buscom. Our main conclusion, again, was that we need a *clear, unambiguous interpretation of the process model*. We can also repeat most of the lessons learnt from the Buscom assessment: the assessment should be carried out in slower pace; one should learn new ways for discussing with management; the assessment process should be planned better; results should be presented both qualitatively and quantitatively, etc.

5 Assessment at Nokia Networks IMN User Interface Team in October 2000

Nokia Networks (NET) IMN organisation develops base stations for the mobile networks. The software developed by the IMN user interfaces team is used for configuring, installing, and maintaining of the base stations.

Context of use process	
Outcomes as defined in KESSU process model	Base practices as defined in HSL model
• Identification of user groups • Description of the characteristics of users • Description of the environment of use • Identification of user accomplishments • Description of user tasks • Identification of user task attributes	• Define the scope of the context of use for the product system • Analyse the tasks and worksystem • Describe the characteristics of the users • Describe the cultural environment/organisational / management regime • Describe the characteristics of any equipment external to the product system and the working environment • Describe the location, workplace equipment and ambient conditions • Analyse the implications of the context of use • Present these issues to project stakeholders for use in the development or operation of the product system

Table 1: Illustration of differences between outcomes and base practices. Example: Context of use process.

5.1 Implementation of the Assessment

Based on the experiences in the previous assessments, the main research driver in this assessment was to develop more unambiguous interpretations of the base practices. We used the *QIU* model (Earthy, 2000), which is the earlier version *HSL* model (Earthy, 2001) — together with the ISO/TR 18529 and ISO 13407 — as references in our interpretation work. The QIU model was recently distributed to a large audience of reviewers, and feedback was desired about it. In addition, we knew that the QIU model was an improved — although more complicated — version of the ISO/TR 18529 model.

This time we had a clear focus in the assessment: one development project. We interviewed the personnel of the company during two weeks' time, having altogether five interviews. In each interview session, there were one or two interviewees. In each interview session, we examined one or two UCD processes.

5.2 Development of New Process and Capability Models

The interpretation of the base practices realised to be quite a challenge. The work led to developing something different than interpretation of base practices. Jointly with the usability experts of NET, we decided to carry out the assessment up to

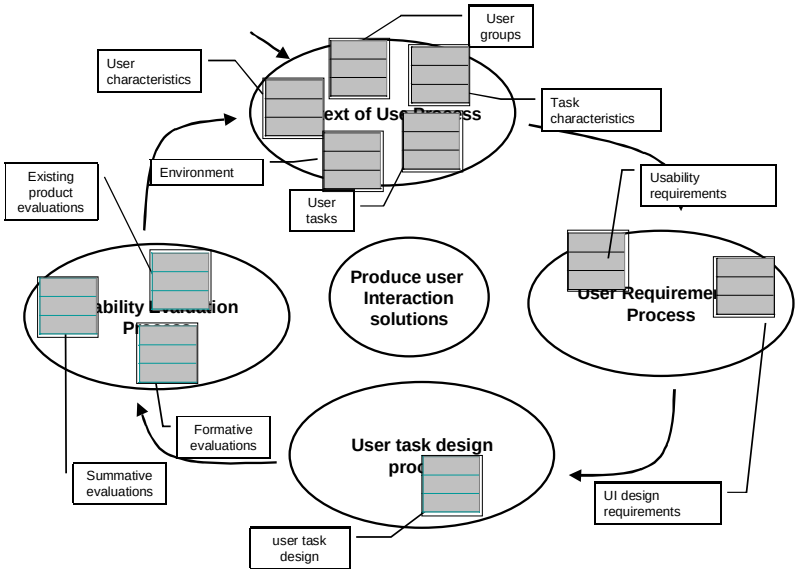

Figure 3: The KESSU Process model.

level 1 based on concrete *outcomes* of processes. In contrast with the outcomes of processes in ISO/TR 18529 (the model defines also outcomes, in addition to base practices), we limited the outcomes to include only concrete deliverables. The outcomes of our *KESSU*[2] Process model and base practices of the HSL model is illustrated in Table 1. As one can see, we have transformed some of the base practices into concrete outcomes. The substance of those base practices that do not produce concrete deliverables are covered by new capability dimensions (see discussion to follow).

Another distinctive feature is that we have split the 'Produce Design Solutions' process of ISO/TR 18529 (and ISO 13407) into two parts. Visually, in its former position there is a 'User Task Design' process. In the centre, there is a new process, 'Produce User Interaction Solutions'. The process model — including the outcomes — is illustrated in Figure 3. We identify five main processes: Context of use process, User requirements process, User tasks design process, Produce user interaction solutions process, and Usability evaluation process.

The reasoning behind splitting the Produce Design Solutions process is that the process is always existent and produces 'full outcome' (the system and user interface) — even in cases where the development process in not user-centred. All the other processes are characteristic to UCD: they provide user-driven information for the design process.

[2]KESSU is the name of our national research project that aims to develop methods for improving user centred-design in development organisations.

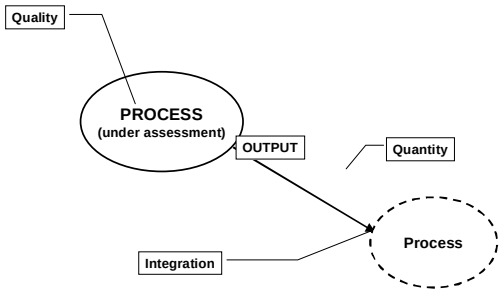

Figure 4: KESSU dimensions of process capability.

The outcome driven assessment led also to new kind of capability dimensions. The capability scale evolved to have three different dimensions as illustrated in Figure 4:

- The *quantity* of outcomes of the process. The more extensively an outcome exists, the higher performance score it gets.
- The *quality* of the outcomes. With this dimension, we examine the quality and validity of the outcomes. For example, we want to make a difference whether an outcome is based on someone's opinions or derived by using recognised user centred methods and techniques.
- The *integration* of outcomes with other processes. The more extensively the outcomes are communicated and incorporated in other relevant processes, the higher rating is given to integration.

The integration aspect has also been addressed in the HSL model. Its solution is, however, different: a process has a specific base practice that addresses integration. An example is shown in Table 1 ("Present these issues to project stakeholders for use in the development or operation of the product system").

The different process and capability models led also to a different way of presenting the results. We present the capability profile in one visual picture, using different symbols to denote the different dimensions, as illustrated in Figure 5.

Altogether, we consider that we have developed new process and capability models. We call the models as *KESSU Process Model* and *KESSU Process Capability Dimensions Model* respectively. They are documented in project reports Jokela (2001b) and Jokela (2001a).

5.3 Feedback from the Assessment

We gathered feedback from the assessment again with questionnaires delivered in the opening briefing, after each interview and finally in the results presentation session. The interviewees found that the assessment approach made a lot of sense. They reported that the interview sessions had pointed out targets for improvement, and

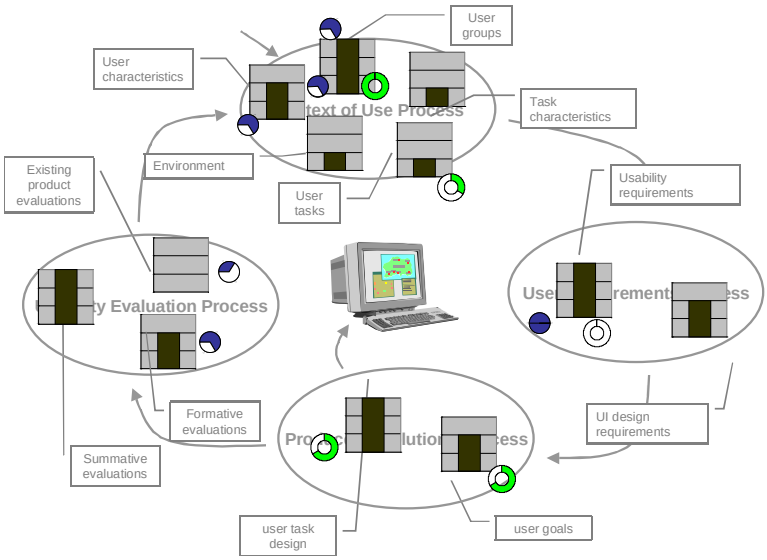

Figure 5: Example of capability profile presentation.

some of the interviewees had received confirmation to their own thoughts on how UCD should be developed.

The feedback from results presentation session was generally good. Especially, the audience reported high motivation for getting training in UCD and for trying UCD approaches in their work. They regarded UCD activities very important in the development.

All those members of the assessment team who had attended the previous assessments found this assessment more successful and sensible than the earlier ones. One illustrative comment from an assistant assessor was: "Interview by interview, the models became clearer. This is the way we should have done from the beginning: to make a clear interpretation of our own about the reference models". The definition of outcomes and assessing through the outcomes was found easier and to give a better picture about the UCD in the organisation than assessment through base practices. Assessors felt that this time they got a good picture about the UCD activities in the organisation.

5.4 Lessons Learnt

A clear feeling after the assessment was that "we want to try this approach again". However, there were also places to improve. The overall role and position of an assessment in the organisational context should be rethought. One organisational problem was, for example, that very few developers attended the opening session of the assessment. In addition, definitions for the levels of rating the capability dimensions should be created, and some terms should be made easier to understand,

for example ‘summative’ and ‘formative’. The assessment report should be quicker to produce.

6 Assessment at Nokia Networks IMN Customer Documentation Team

This assessment had a different object domain than the earlier ones: the object of the assessment was the development process of customer documentation. Three key persons from the customer documentation and a two-person assessment team participated the assessment.

The customer documentation group has experience about UCD for some years. For example, they have organised user evaluations of user manuals in laboratories and in the field. There is clear enthusiasm towards UCD in the group. The group is further developing their documentation processes, and was motivated to get feedback from an external evaluator.

The assessment had one significant difference compared with the earlier ones. It was decided to carry it out in half a day workshop. The reason for this was practical: there was no time for an assessment with multiple interviews.

6.1 Implementation of the Assessment

We used the process and capability models developed in the previous assessment. Our plan for the assessment was to follow a cycle that is used in typical assessments: interview, wrap-up of the findings by the assessment team, and agree on the results with the interviewees. At this time, we just hoped to cover all the relevant processes in a time that normally is reserved for one long interview.

In the interview part, we asked the project team to describe the activities of the development process of the user manual and quick reference guide. We did not describe our reference model beforehand but used it as our reference in the interview.

In the results session, the findings were contrasted against the process model. The lead assessor explained the model step by step, and interpreted the information received in the interview session against the model. We agreed on the areas where the organisation had strengths and weaknesses. The session was finished in about 1.5 hours. Some refinements and clarifications were made on the process model during the discussion.

6.2 Feedback from the Assessment

One immediate feedback came actually when we finished the workshop. One of the participants said that she already had started to challenge herself on how to implement improvements in one specific area. We asked feedback from the project team through email. Comments were received from the team on the same afternoon. They reported for example:

- “The assessment pointed out many issues to consider in the following documentation projects.”
- “Now we know that task analysis is important. We also need to work on usability requirements.”

- "We found that your model worked well also in our domain."

The assessment team found the assessment as a positive experience. We succeeded in getting a credible picture of the development process in a short time. We had felt that there was a positive atmosphere in the workshop.

6.3 Lessons Learnt

This assessment confirmed that we would experiment the revised process and capability models in the forthcoming assessments. The definitions of processes seem to make UCD concrete to the organisation. The capability dimensions point out the areas of improvement, and give a means to discuss the project at an appropriate level of abstraction.

The specific implication of this assessment is to try the workshop approach again. It is efficient — one workshop instead of a series of interviews. Moreover, a workshop may be a solution to one problem that we have faced: people being in different positions in assessments (those who are interviewed, and those who are not).

7 Discussion

Our target was to learn how to perform effective usability capability assessments through experiments in industrial settings. For that, we carried out four experimental usability capability assessments.

In the beginning of the research, we used an ISO 15504 style process assessment approach with one capability dimension, rating capability through base practices, and an assessment method with a number of interviews. During the assessments, we found that in our context it sensible to focus on the performance — not management — of UCD activities. As a result, we developed an assessment approach with an outcome-driven process model, three-dimensional capability model, and an implementation the assessment as a workshop.

Our process and capability models seem to make the assessment and UCD more understandable to the audience and easier for the assessors in our context. The workshop type of assessment makes an assessment efficient and spreads UCD knowledge to larger part of the staff than interviews of a limited number of people.

We, however, want to emphasise that our assessment approach is a complementary one to the traditional process assessment. Compared with ISO 15504 assessment, we can say that we examine the UCD activities thoroughly 'below the level 1 of capability'. Traditional process assessment should be used in contexts where it is applicable to examine the management of UCD activities at higher levels of capability.

7.1 Contrasting the Results with Experiences of the TRUMP Project

Bevan & Earthy (2001) report about recent assessment case studies carried out in the European TRUMP project. They carried out assessments in two organisations: at Inland Revenue (IR) at UK and at Israel Aircraft Industries (IAI). In both assessments, they used the ISO/TR 18529 as a reference. The IR assessment was a 'full'

process assessment with twelve interviews while the assessment at IAI was a light assessment, a one-day workshop. Bevan & Earthy report that both assessments were successful.

One can regard the assessment at IR and our assessment at Teamware methodologically similar: traditional process assessment based on the same process model. The assessment at IR, however, was seemingly a more successful than the one at Teamware. One potential explanation may be the fact that the lead assessor in the IR assessment was the main author of the ISO/TR 18529 model. The assessment team at IR was also probably more experienced in process assessment than we are. In addition, we find that there are many other factors that potentially affect the success of an assessment. These factors include: the business and products of the company; general situation in the organisation (e.g. whether there has been recent big changes); the culture of the organisation; the tradition of carrying out development projects and improvement programs; the position of the sponsor of the assessment; how the assessment is planned, organised and conducted; how the project(s) to be assessed are selected; how successful are the presentations; the way the processes and base practices are communicated in the interviews; the characters and attitudes of individuals; the position of usability persons in the organisation etc. Not only the assessment approach but these kinds of non-technical factors may have a great impact to the success of the assessment.

The assessment at IAI resembles our last workshop at NET. They both were successful, too. The reference models were a bit different — IAI was assessed by using base practices while we used KESSU process model as reference. One may ask whether we would have succeeded with base practices, too. It may be possible. We find 'outcome based assessment' working probably because it makes things concrete — the outcomes are deliverables. Using 'concrete' base practices might work as well.

Bevan & Earthy's report that the success of the improvement efforts in the TRUMP trials has been very good. We have not yet carried out subsequent assessments to monitor our success. We assume that our success has not generally been at the same level. On the other hand, we find that the success of improvement actions does not depend on the success of an assessment only. Success in improvements depends on many factors — as do the assessments.

7.2 Limitations

There are some limitations to be considered when making decisive conclusions about our assessments experiences. First, the organisational situations may be very different, and one assessment approach that is suitable for one organisation may not be the best choice for another one. Therefore, comparing the success of assessments based on the feedback from the organisations assessed is problematic. On the other hand, we find as an advantage in our research that the assessment team was almost the same in all assessments. The assessment team was able to compare the different assessments.

Second, our goal for the assessments was to give a good basis for improvement actions in UCD. From this viewpoint, issues such as spreading knowledge about UCD to the staff and getting them committed to UCD improvement actions are important.

A different target for assessment could be for example to get exact ratings of usability capability for selection of contractors. This was not our goal. Another viewpoint that we excluded is standardisation that has been one driver of the ISO/TR 18529 and HSL models.

Third, each assessment is a different instance, even if the same approach is used. Assessment is a very human process, and its success may depend on many human issues — both the organisation assessed and the composition of the assessment team have influence on it. Specifically, an assessment necessarily has 'a look' of the lead assessor.

Fourth, one limitation is related to the assessment with the ISO/TR 18529 and HSL models. The assessments were very much based on the interpretations, experience and style of the lead assessors. The interpretation of the ISO/TR 18529 model is based on documentation. Some other person may have conducted the assessments in a different way.

Fifth, a specific feature in our assessments is that all organisations represented geographically limited industrial settings.

7.3 Implications for Next Assessments

We started with assessments with a number of interviews using a traditional process assessment approach, and finished with an assessment that was half a day workshop. What are our choices for the next assessment? In our environment, we most probably go for a workshop. We were able to get a clear picture of the position of UCD in the organisation, and to analyse and communicate the results in a very short time in the last assessment at NET. Next time, however, we will assign more time — probably one day — for the workshop.

We will continue with the KESSU process model and capability scale (probably with refinements). We find that the outcome-driven process model makes discussions concrete, and the three-dimensional capability scale makes possible to have an appropriate level of abstraction in the assessment.

For those who have not carried out assessments before, our main advice is that one should understand that an assessment is a research effort, no matter which model the assessment is based on. There exist very few reported experiences on assessments. No assessment approach — inclusive the one that we use — is so matured that one can totally rely on its validity.

7.4 New Research Topics

As said, each new assessment should be considered a research activity. A researcher should try to get access for following assessments and for gathering feedback from them. There is definitely space for improvements both at the model and in the assessment method (steps of assessment) levels. We will carry out further assessments, and regard each of them also as a research effort. — Actually, we just finished another assessment that was implemented as a workshop. The most important findings of this assessment are that one should understand the organisational improvement context before planning an assessment and take carefully into accout the human aspects. We will report the findings of this assessment in a forthcoming paper.

The next step in our research is of very constructive nature: to document the assessment approach as a handbook. We assume that the creation of such a document is not a one-time effort but it will be revised after trials. Another artefact we plan to develop is a template to make the assessment efficient. We hope to have an online documentation in the workshops, and deliver the results immediately.

One interesting challenge is how to reliability verify the success of an assessment. We have gathered a lot of feedback in our assessments. However, it still is difficult to make definite conclusions. For example, we find that the assessments should be also training occasions where the understandability of models and results is important. Some others may disagree with the importance of this criterion.

Acknowledgement

There are a number of people who have contributed this work. We wish thank the personnel of the partner companies, especially Erkki Jylhä-Ollila, Risto Rauhala and Timo Salmu from Buscom; Pirkko Jokela and Eija Suikola from Teamware; Pekka Suhonen, Marjo Favorin, Tuula Stromberg, Jouko Salo, Johanna Tapio and Tanja Petrell from Nokia Networks; and Petri Hyyppä from Nokia Mobile Phones. We also wish to express thanks to the other project members of the KESSU project: Pekka Abrahamsson, Mikko Jämsä, Tero Posio, Tonja Molin-Juustila, Mikael Johnson, Samuli Saukkonen, and Kari Kuutti; and to the national funding organisation TEKES.

References

Anderson, R. (2000), "Organisational Limits to HCI. Conversations with Don Norman and Janice Rohn", *Interactions* **7**(3), 36–60.

Bevan, N. & Earthy, J. (2001), Usability Process Improvement and Maturity Assessment, *in* J. Vanderdonckt, A. Blandford & A. Derycke (eds.), *Proceedings of IHM-HCI'2001, Joined Conference on Human–Computer Interaction: Volume 2*, Cépaduès-Editions.

Bloomer, S. & Wolf, S. (1999), Successful Strategies for Selling Usability into Organizations, *in* M. W. Altom & M. G. Williams (eds.), *Companion Proceedings of CHI'99: Human Factors in Computing Systems (CHI'99 Conference Companion)*, ACM Press, pp.114–5.

Earthy, J. (1999), Usability Maturity Model: Processes, Project Report, Lloyd's Register of Shipping, London, UK.

Earthy, J. (2000), Quality In Use: Processes and Their Integration — Part 2, Assessment Model, Project Report, Lloyd's Register of Shipping, London, UK.

Earthy, J. (2001), Ergonomics — Human System Interface — Human-system Life Cycle Processes Proposal, Project Report, Lloyd's Register of Shipping, London, UK.

ISO (1998a), "Software Process Assessment — Part 2: A Reference Model for Processes and Process Capability", Project Report. International Organisation for Standardization, Genève, Switzerland.

ISO (1998b), "Software Process Assessment — Part 7: Guide for Use in Process Improvement", Project Report. International Organisation for Standardization, Genève, Switzerland.

ISO (1999), "ISO 13407 International Standard. Human-centred Design Processes for Interactive Systems". International Organization for Standardization, Genève, Switzerland.

ISO (2000), "Human-centred Lifecycle Process Descriptions", Project Report. International Organisation for Standardization, Genève, Switzerland.

Jokela, T. (2001a), KESSU Process Capability Dimensions, v0.1, Project Report, Oulu University, Oulu, Finland.

Jokela, T. (2001b), KESSU Process Model, v0.2, Project Report, Oulu University, Oulu, Finland.

Kuutti, K., Jokela, T., Nieminen, M. & Jokela, P. (1998), Assessing Human-centred Design Processes in Product Development by Using the INUSE Maturity Model, *in* S. Nishida & K. Inoue (eds.), *Proceedings of the 7th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man–Machine Systems (MMS'98)*, IFAC, pp.89–94.

Kuvaja, P., Similä, J., Kranik, L., Bicego, A., Saukkonen, S. & Koch, G. (1994), *Software Process Assessment and Improvement — The BOOTSTRAP Approach*, Blackwell.

McFeeley, B. (1996), IDEAL SM: A User's Guide for Software Process Improvement, Project Report CMU/SEI-96-HB-001, Software Engineering Institute, Pittsburgh, USA.

Paulk, M. C., Weber, C. V., Curtis, B. & Chrissis, M. B. (eds.) (1995), *The Capability Maturity Model: Guidelines for Improving the Software Process*, Addison–Wesley.

Rosenbaum, S. (1999), What Makes Strategic Usability Fail? Lessons Learned from the Field, *in* M. W. Altom & M. G. Williams (eds.), *Companion Proceedings of CHI'99: Human Factors in Computing Systems (CHI'99 Conference Companion)*, ACM Press, pp.93–4.

Rosenbaum, S., Rohn, J. & Humburg, J. (2000), A Toolkit for Strategic Usability: Results from Workshops, Panels, and Surveys, *in* T. Turner, G. Szwillus, M. Czerwinski & F. Paterno' (eds.), *Proceedings of the CHI2000 Conference on Human Factors in Computing Systems, CHI Letters* **2**(1), ACM Press, pp.337–344.

Author Index

Keyword Index