



Mountain meadow, P. Prusinkiewicz, R. Mech, M. Pharr (1998)

Context: An Active Choice in Usability Work

The concern for usability in the Danish software industry has become more visible because of usability labs or usability groups within software companies. In the BIDI (Brugbarhedsarbejde I Dansk Industri, in English: Usability work in Danish industry) project, we are collaborating with Bang & Olufsen, Danfoss, and Kommunedata, the first three Danish companies to have usability lab facilities.



Suzanne Bodker
Department of
Computer Science
University of Aarhus
Ny Munkegade,
bldg. 540
DK-8000 Århus C
+45 8942 3256
bodker@daimi.aau.dk

Kim Halskov Madsen
Department of
Information and
Media Science
University of Aarhus
Niels Juelsgade 84
DK-8200 Århus N
+45 8942 1958
halskov@imv.aau.dk

Bang & Olufsen is a manufacturer of integrated video and audio products that are linked in a small network and that allow shared service (e.g., a person from one room can select and listen to a track from a CD player that is placed in another room). The prime characteristic of these products is that they are part of people's everyday environment.

Danfoss is a manufacturer of mechatronic products such as flow meters, temperature sensors, and controls. These products are used in a diversity of settings that range from private homes, apartment buildings, supermarkets, and district heating stations to waste water treatment plants. The products are often only a small part of a complex piece of technology.

Kommunedata is the principle supplier of administrative systems for local city administrations in Denmark. They develop computer systems with a standard keyboard and screen for a diversity of work settings (hospitals, dental clinics, municipal offices, etc).

The BIDI project is an action-oriented research project, based on our own work as well as the work of others [3, 4, 8], that aims to develop the work practices of usability professionals. Our theoretical and methodological platforms include participatory analysis and design, activity theory, and ethnographical method. The goals of the project are (a) to bring test situations closer to the nature of the future situation of use, (b) to explore new kinds of user participation, (c) to explore how learning takes place and how new patterns of use evolve, (d) to investigate usability issues related to new kinds of products beyond a conventional interface with screen, mouse, and standard keyboard, (e) to support self-reflection among usability professionals, (f) and to explore how field studies and lab tests can mutually support each other.

In this article, we focus on how to understand *use context* and, particularly in various situations, how to create the appropriate context for usability work (*evaluation context*). Because the full-fledged use situation in general is extremely complex, a reduction of complexity with respect to certain purposes is

necessary, but how to make such a reduction is not obvious. In order to support the choice of work methods whereby the issue of context is treated in a careful manner, we (a) provide a checklist of questions to help focus on the evaluation situation; (b) discuss a number of particularly important aspects of context: work materials, artifacts, time, cooperation, location and physical space, and organizational setting; and (c) present a repertoire of ways of creating evaluation context.

Choosing Method

Choosing methods for usability testing is not a matter of testing in the lab on the one hand versus studying the completed design in the field on the other. Furthermore, such choices depend on each other, because what may seem like an option under one set of conditions may no longer be viable after other usability evaluations have been carried out. The choice of methods at a particular point in time depends on a number of characteristics of the evaluation situation, including (a) the purpose of the evaluation activity, (b) the knowledge of the context, (c) the access to the workplace, (d) the resources available for the evaluation, and (e) the availability of prototypes or other design artifacts.

To support the choice of work methods, whereby the issue of context is treated in a careful manner, we provide a checklist to help focus on the evaluation situation.

1. What is the purpose of the evaluation situation?
 - ✗ Understanding current practice and context.
 - ✗ Understanding the future practice and context.
 - ✗ Generating alternative ideas for a particular kind of artifact.
 - ✗ Getting a proof of existence of a particular artifact.
 - ✗ Testing a particular solution.
 - ✗ Showing which contextual issues are of concern (work materials, artifacts, time and timing, cooperation, location and physical space, organizational setting).
2. Which aspects of the context are known and how well?

- ✗ Artifacts.
 - ✗ Time and timing.
 - ✗ Cooperation and communication.
 - ✗ Location and physical space.
 - ✗ Organizational setting.
3. Which access can be obtained to the workplace?
 - ✗ Access to the work site.
 - ✗ Access to the users, at their work setting or outside it.
 - ✗ Established active cooperation with users.
 4. Which resources are available for the evaluation and which are the scarce resources?
 - ✗ People.
 - ✗ Hours.
 - ✗ Equipment.
 - ✗ Financial resources.
 5. How stable is the artifact that is to be evaluated?
 - ✗ Is it mainly at the level of a concept?
 - ✗ Is it a paper prototype?
 - ✗ Is it a running horizontal or vertical prototype?
 - ✗ Is it a beta version or a full application?

In the following, we look at a couple of examples from our project to illustrate how the questions may support our choice of work methods.

Danfoss Case

Danfoss has carried out a field study of the work at the Sønderborg combined district heating and power plant. It was carried out in several rounds encompassing three to eight people who spend several days at the plant, and the study investigates coordination and

communication and physical space and location. This field study is seen as an initial activity of “The Smart Window” project in which Danfoss is exploring portable user interfaces of the future. The field study has been the starting point for an ongoing interaction with the workers of the plant, including discussing a video from the plant and discussing early mock-ups with the workers both in workshop settings and in their real work environment.

In connection with a workshop, we discussed how to evaluate a prototype, which is not in a state in which it would fully replace the current tools and procedures, in order to understand if such a solution is feasible and to get further ideas for the design of the human–computer interaction. It is evident that Danfoss wants to go back to the users at the Sønderborg plant to have them evaluate the prototype testing of a particular solution with respect to real materials and artifacts, as well as cooperation.

The relations with the workers at the plant are obviously fairly well established. At the same time, Danfoss has already taken a lot of time from the daily work time of these workers. This introduced a discussion of testing with limited personnel resources in a situation in which there is real work to be done: It would not be possible to stop the real work of the plant for a day nor to do a simulation using the prototype instead. It would also not be possible to run the plant (even partially) using the prototype. The key characteristics of the situation are summarized in Figure 1.

Danfoss situation	
Purpose	Testing a particular solution with respect to real materials and artifacts, as well as cooperation
Context known	Cooperation, physical space and location
Access to workplace	Well-established relations to users
Resources	Enough money, but limited time for users and for the actual testing
Artifact	Horizontal prototype that does not interact with other artifacts in the environment

Figure 1. The situation at Danfoss in “The Smart Window” project.

We discussed two possibilities: (a) a kind of shadowing approach in which a person using the prototype follows and redoes the procedures undertaken by the person doing a real task and (b) a simulation of real tasks based on a reconstruction of the tasks from videotapes.

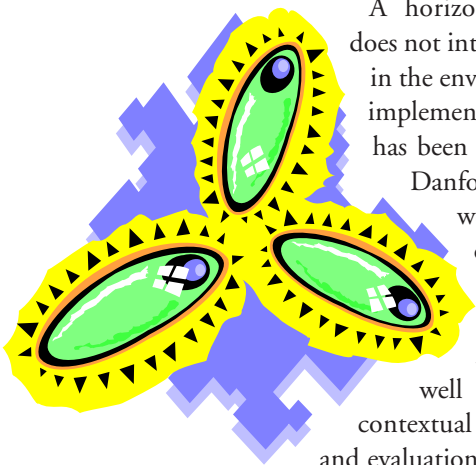
In both cases, it would be possible to do the evaluation without interfering with actual work of the plant if Danfoss could persuade (and pay) somebody off-duty to do the evaluation.

The conclusion was that the shadowing is more suitable for what may be seen as routine situations, such as the daily rounds in which the instruments, etc., of the plant are controlled. It would be harder to use this approach for problem situations such as the handling of serious alarms from the plant.

To appreciate fully the argument for why problem situations are particularly challenging, another discussion needs introduction: the completeness of the prototype and test data. In order to prepare for testing surprise situations such as real alarms, it is necessary to have a complete prototype and set of test data. On the other hand, if we know with which alarm we are dealing, we may limit the test data to what is necessary to handle this alarm. To delimit the evaluation to one or two alarms that are already known to the designers/evaluators may seem to be a heavy limitation. However, as most alarms are rather standard anyway, it seems less limiting in this particular case.

A horizontal prototype, which does not interact with other artifacts in the environment, has later been implemented on a Newton and has been shown to clients within

Danfoss. The prototype was well received, and based on this, Danfoss will continue to work on portable interfaces in a number of projects, as well as to experiment with contextual approaches to design and evaluation.



Kommunedata Case

In one of Kommunedata's recent field studies, they wanted to go beyond using field studies as an element in identifying test cases. The purpose of field studies was primarily methodological: The aims were to develop approaches that would further include time, cooperation, and physical and organizational setting in situations in which access to the actual use situations is limited.

In an earlier round, the lab tested a nursing-care-plan system for hospitals and identified several problems. They had further done a field evaluation of a beta version of the system in which the focus was to see if the problems identified earlier still occurred after a period of use and if new problems had emerged. During the field evaluation, they gained some knowledge about time aspects, cooperation, and the physical and organizational setting.

In this field study, Kommunedata wanted to get beyond this narrow problem focus while concurrently making use of prior experiences with the field setting as well as the product. Two usability testers visited two particular wards at two hospitals, one day each place, and observed the use of the system. The usability people ended up not only taking an observational role but also discussing the work practice with the nurses.

At a series of project meetings in BIDI, the field studies were discussed, and a number of aspects of conducting them were considered. Because patients cannot be filmed, following staff on rounds was considered as an alternative. Another point made during the discussion of the field studies was the fact that each day at the ward seems to have a particular rhythm with particular kinds of main activities: when a nurse begins a shift, ends a shift, etc. The key characteristics of the situation at Kommunedata are summarized in Figure 2.

In discussing how knowledge about use might inform the work in the conventional usability work at Kommunedata, the BIDI project members discussed how to deal with the fact that nurses are continuously interrupted from their use of the system by calls



Kommunedata situation	
Purpose	Primarily methodological: The aims were to develop approaches that would further include time, cooperation, physical and organizational setting
Context known	Some awareness of time, cooperation, physical and organizational setting
Access to workplace	Limited access to the work site
Resources	Usability group has little resources. Normally at most two testers are assigned to a task.
Artifact	Beta version

Figure 2. The situation at Kommunedata.

from patients, doctors, etc. One of the possibilities discussed was a simulated work situation in which one would create “situation cards” [6] stating various kinds of interruptions and then use these situation cards in random order to simulate interruptions during lab test.

For planning future test sessions, an obvious possibility would be to reflect the variations in daily patterns in different kinds of test scenarios. This is both a matter of recognizing the differences between the patterns and creating the necessary realistic time pressure for the evaluation situations.

Aspects of Context and How They Might Be Dealt With

In the previous section, we suggested a checklist to help focus on the evaluation situation. In this section, we look further into one of the key parts of the checklist, the context itself. The concern for the use context on the one hand and the evaluation context on the other is a key issue. However, use context is immensely complex, and full-fledged field situations are far too complex to come into action for every usability evaluation situation we choose to set up. Thus, identifying the context for a specific usability situation is very important: Choosing which aspects of use context must be brought into the test situation, which may matter less for the particular situation, and which particular context needs to be staged to make the usability situ-

ation work.

Although the complexity of context can obviously not be reduced to a number of independent dimensions, we do, however, argue that it is possible to identify a number of aspects of context to be considered when choosing an evaluation method. Organized around various aspects of context, we present a repertoire of various ways of studying context and staging an appropriate evaluation context.

Work Materials and Other Artifacts

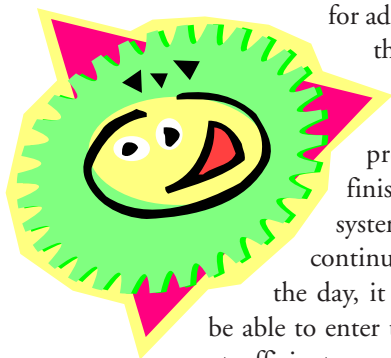
One needs to be concerned with investigating the work situation so as to retrieve realistic test data to be put into prototypes, for example, to be tested. A variety of ethnographic and other methods apply for the uncovering of such test data [2], but as the Danfoss example shows, the next step is far less obvious: When testing a prototype, it is often impossible to base this on all possible realistic test data and situations, so which test data to include must be aligned with which situations one wants to test and the stability of a prototype.

Second, the concern must be for the extent to which users have their everyday materials and tools available alongside the product being tested. This concern is a matter of both adding realism to the test situation and compensating for the lack of realistic test data in a prototype. Bødker and Grønbæk [5] describe a case in which a municipal planning office’s participating case workers each had his or her

frame task to work on in the sessions. These tasks were representative of the work done by the caseworkers, and the aims were to create prototypes that would simulate support for these selected work tasks. In the sessions, the caseworkers brought the appropriate work materials of the task (e.g., urban plans, maps, etc). In some of the examples, the caseworker primarily demonstrated his or her current role when going through a typical work task, the frame task, even though they were detached from their normal performance of the task. This was made possible because of the work material, such as maps, letters, urban plans, and other stuff, brought into the sessions.

Time

Time is a difficult matter in usability evaluation because “time” in the use context often significantly differs from “time” in the evaluation context. First, real-use situations are often sporadic-use situations, such as people’s use of Bang & Olufsen stereos at home. Second, use is often interrupted by a variety of events, such as the nurses being paged for an emergency or the power plant maintenance being interrupted by an alarm. We claim that interruptions are the rule rather than an exception in most work situations, and an evaluation of use that does not consider this rather unpredictable phenomenon is a problem. A further issue of time and timing is that of variation of work over time: In the hospital ward, the use situations are different depending on shift. For example, the night nurses have more time



for administrative work than the day shift, and they are less likely to be interrupted. The pressure on nurses to finish by reporting to the system increases as the day continues, and at the end of the day, it is simply necessary to be able to enter the information in the most efficient way in order to pass the information to the next shift.

In a study of a law office [1], the sporadic use of a document file cabinet was handled by putting up a camera that the attorney would

turn on when he used the file cabinet for his own purposes or in response to colleagues’ requests. Another approach was taken in one of the field studies of Bang & Olufsen in which the focus was on how some of their products were used as part of people’s everyday life in their homes. Because of the sporadic use and because doing pure observations in a person’s private home would not be appropriate, Bang & Olufsen did some interviews in homes that included asking people what they did the last time they used the audio and video devices combined with informal simulations in which people were asked to act out typical situations of use, such as what they do when they want to listen to the radio in the morning.

As suggested in the discussion of Kommunedata’s field studies, interruptions may be simulated using “situation cards” during lab tests. The fact that each day at the ward seems to have a particular rhythm with particular kinds of main activities could be reflected in different kinds of test cases and scenarios. In order to further the realism of time pressure put on people, it is worth considering how to put a realistic time pressure on the evaluation situation so that it reflects the time pressure in the real situation. A realistic solution to this would be to ask a user to enter a specific, realistic set of observations while the test is being performed.

Cooperation

In the Kommunedata field study, it turned out that a lot of interaction took place around the use of the system. Thus, the use of the system emerged in a cooperation between users, an aspect that Kommunedata had not been aware of in the previous lab tests. It would be difficult to study such emergent features in a simulation or lab setting, because a fundamental study of the emergence of use is a much more long-term agenda that would probably require a more dramatic change of work method. What we are after would instead be to be aware of these sorts of things happening and to allow for some appropriate kind of cooperation in the evaluation, when that is a key issue in the use

context. A first obvious step is evaluating with more than one user. Second, there are various possibilities for making it possible for people who know the artifact and people who don't, though they know the work domain, to collaborate or to use the same test group over a longer period of time, testing various versions of the artifact.

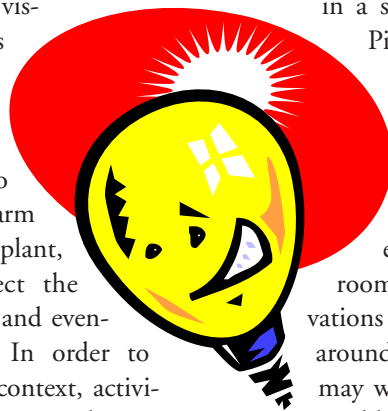
A more explicit kind of cooperation took place in the Sønderborg plant. In order for somebody in the plant to turn off a pump, for example, he or she needed to cooperate with the control room, where it was visible which valves and pumps were feeding the particular pipe. Debugging of alarms was done in cooperation between the control room operator, who would be able to read the alarm data, and the person in the plant, who would physically inspect the devices, find causes of alarms, and eventually handle the problems. In order to understand this aspect of use context, activities were videotaped both in the control room and out in the plant. These tapes are a valuable source for establishing the two evaluation situations discussed earlier, both for understanding what is likely to happen in the shadowing of a real round and when preparing data, etc., for simulation of a particular alarm.

Location and Physical Space

In the Kommunedata case, nurses used the system in a room separate from the actual ward, and at Danfoss, we were dealing with a case in which people were physically distributed, in which normally the overview of the process is located in the control room, and in which there is a heavy communication between the control room and the personnel moving around the plant. In both cases, the issue of mobility was brought up: Users have to move to particular locations in order to work with technology—in the plant case, the control room, and in the case of the hospital ward, to the room designated for computers. In establishing the evaluation context, it is very important to be aware of the difference between contexts in which technology and

other important artifacts and materials are secluded in particular physical locations and contexts in which this is not the case.

In both cases, there are obvious reasons for pondering the issue of what would be different if users had access to the information when they were on the spot where they need to act, whether this is by the flow meter in the plant or by the hospital bed. There is a variety of technical, work organizational, and other questions attached to exploring such scenarios, on top of the location questions: If not in a separate room, then where?



Picking up the hospital ward example, it is possible to have nurses carry a portable computer on rounds. Other options are to place computers next to every bed or at least in every room. However, given the observations of how nurses cooperate around the use of the system, one may wonder whether it would be possible to maintain such cooperation with a changed physical layout?

Kjær and Madsen [7] have used what they call *blueprint mapping* as a technique to address the relation between technology, work activities, and physical space. The idea is to start out from a regular blueprint of the work site and identify how various kinds of technologies make information available at different locations for the various work activities.

Organizational Setting

In Bødker [3], it is illustrated how users' evaluation of a system is heavily dependent on the organizational situation: Managerial users in the case had primarily a need for overview statistics, whereas the secretaries who were actually entering data and extracting reports had an entirely different understanding of the system. A third group, case workers who were planning their own work and accounting from it using the same system, had yet another view. In the Kommunedata case, the first step was to recognize that there were different sorts of users, some of whom were only indirect users of the system, such as doctors: The field



METHODS & TOOLS COLUMN EDITORS

Michael Muller

mullerm@acm.org

Finn Kensing

Computer Science

Roskilde University

P.O. Box 260

DK-4000 Roskilde

Denmark

+45-4675-7781-2548

fax: +45-4675-4201

kensing@dat.ruc.dk



studies revealed that the care-plan system had more kinds of users than initially anticipated by the usability workers and designers.

Understanding the difference between the three shifts and their use of the system is another such example. In the Danfoss case, one observation is that although one person has the responsibility for monitoring the plant from the control room and another is “walking the plant,” the division of work is such that they take turns every second day. This means that these two persons know and are able to act in both positions.

For a usability evaluation, it will matter with whom one may choose to work. The evaluation that one would get from a secretary, a case worker, or a managerial user is likely to differ, which means that the choice is important. It is also important to establish the organizational setting and role of the people involved in evaluation so that they know in which capability they are acting, which cases have discretionary power, and which do not and also so that they know with which other capabilities they are working, etc.

We propose to seek inspiration from organizational games (e.g., see [6]) and scenarios in order to establish the organizational context. This is not to establish artificial or nonexistent organizational roles but rather to be specific as to whom the users are reporting, in what capabilities they are acting, and which decisions they can make themselves. This would make it possible to make an evaluation that is more grounded in the actual organizational context and actually to “test” the assumptions that are made about the organizational context in the evaluation.

In this section, we presented a repertoire of various ways of studying context and staging an appropriate evaluation context:

- ◆ Use realistic test data

- ◆ Align test data with the aspect to be tested
- ◆ Align test data with the stability of the prototype
- ◆ Add realism to the evaluation by having everyday materials and tools available alongside the product being tested
- ◆ Have users operating a camera as a way of capturing sporadic use
- ◆ Ask users “what they did last time they used...” as a way of staging otherwise only sporadic use
- ◆ Simulate interruptions by using “situation cards”
- ◆ Involve more than one user at a time
- ◆ Use videotapes as sources for establishing the evaluation context
- ◆ Use blueprint mapping as a technique to address the relation between technology, work activities, and physical space
- ◆ Use organizational games and scenarios in order to establish the organizational context

Choosing Work Methods in Context

Context is immensely complex, and we have drawn to attention a number of specific aspects of context that we consider to be of particular importance. We have further discussed how it is important to establish a proper evaluation context. In order to support the choice of evaluation and work methods, we have provided a checklist of questions to help focus on the evaluation situation that one would want, the issues that make it worth while, and the resources needed. The list of methods to choose from or to develop for the particular purpose are, however, nearly endless. Instead of listing and characterizing these, we have chosen to give examples. These examples would hopefully help the reader reflect on his or her own work practice and imagine new ways of working. In specific situations, there is not one choice that will satisfy all needs and help focus on all relevant issues. There are trade-offs that can only be dealt with in and for the particular situation, and even then, one choice may lead

to even more open questions.


Acknowledgment

The BIDI project is sponsored by CIT through project #23. We would like to thank the usability groups from the three Danish companies.

References

1. Blomberg, J., Suchman, L., and Trigg, R. Reflections on a work-oriented design project. In Trigg, R., Anderson, S., and Dykstra-Erickson, E. (Eds.), *PDC'94 Proceedings*. 1994, pp. 99–109.
2. Blyth, S., Rouncefield, M., and Hughes, J. Never Mind the Ethno' Stuff, What Does All This Mean and What Do We Do Now: Ethnography in the Commercial World. *Interactions* 4, 3(1997), 38–47.
3. Bødker, S. *Through the Interface—A Human Activity Approach to User Interface Design*. Hillsdale, NJ: Lawrence Erlbaum Associates.
4. Bødker, S. Applying activity theory to video analysis:

How to make sense of video data in HCI. In Nardi, B. (Ed.), *Context and Consciousness. Activity Theory and Human Computer Interaction*. Cambridge, MA: MIT press, pp. 147–174.

5. Bødker, S. & Grønbæk, K. (1991). Cooperative Prototyping: Users and Designers in Mutual Activity. *International Journal of Man-Machine Studies*, 34, Special Issue on CSCW (453–478).
6. Ehn, P., and Sjögren, D. From systems descriptions to scripts for action. In Greenbaum, J., and Kyng, M. (Eds.), *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, NJ: Lawrence Erlbaum, pp. 241–168.
7. Kjær, A., and Madsen, K. H. The dependencies between work activities, technical artifacts, space, and work organisation: Flexibility issues. *Scandinavian Journal of Information Systems* 7, 2, pp. 23–44.
8. Madsen, K. H. Initiative in participatory design. In Blomberg, J., Kensing, F., and Dykstra-Erickson. PDC '96, *Proceedings of the Participatory Design Conference*. Palo Alto, CA: Computer Professionals for Social Responsibility, pp. 223–230. 

PERMISSION TO MAKE DIGITAL OR
HARD COPIES OF ALL OR PART OF
THIS WORK FOR PERSONAL OR CLASS-
ROOM USE IS GRANTED WITHOUT FEE
PROVIDED THAT COPIES ARE NOT
MADE OR DISTRIBUTED FOR PROFIT
OR COMMERCIAL ADVANTAGE AND
THAT COPIES BEAR THIS NOTICE AND
THE FULL CITATION ON THE FIRST
PAGE. TO COPY OTHERWISE, TO
REUBLISH, TO POST ON SERVERS OR
TO REDISTRIBUTE TO LISTS, REQUIRES
PRIOR SPECIFIC PERMISSION AND/OR A
FEE. © ACM 1072-5220/98/0700
\$5.00