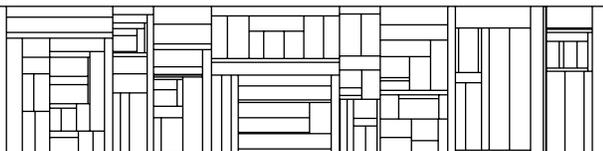# Re-framing research in human-computer interaction from the point-of-view of activity theory?

Susanne Bødker

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF AARHUS

Ny Munkegade, Bldg. 540
DK-8000 Aarhus C, Denmark

Re-framing research in human-computer interaction from the point-of-view of activity theory?

Susanne Bødker

Introduction

Within the area of human-computer interaction cognitive science is today the leading approach. Yet its position is threaded philosophically by for instance T. Winograd & F. Flores (86), L. Suchman (87), and H. & S. Dreyfus (86). Their critique is part of an evolving theoretical approach which is considering human activity, including design or use of computer applications, not as primarily characterized by rationality, planning, and reflection, but by practice and our ability to act in situations, which are more or less familiar to us, where reflection is something secondary or 'post factum'. Even from within the field of human-computer interaction the position of cognitive science is under debate because of a growing concern for practical use of the theory. In Bannon & Bødker (1991) it is pointed out, though, that most attempts to re-frame human-computer interaction from within the field are minor revisions to the traditional theoretical basis, still without much to say about real-life computer applications. In this paper I shall present and discuss an alternative approach based on activity theory (Bødker 1990, 1989, in preparation).

From the point of view of human-computer interaction, the work of Leontjew, Vygotsky et al. has the advantage that it lets us study practice in relation to the material conditions shared by a group of human beings, and that it situates praxis historically. The following summarize the theoretical anchor points:
• Activity is *mediated*. The mediation is essential when understanding artifacts. That artifacts mediate use means that we are normally not aware of them as such in use, they are transparent to us. Our attention is with the materials and products even though many of our experiences with these are made through the artifact. In this perspective we cannot study computer applications as things, we need to look at how they mediate *use*. Furthermore, artifacts are not just means for individuals, they also carry with them certain ways of sharing and dividing work. They are given meaning through their incorporation into social praxis.
• Artifacts are seen as *historical* devices which reflect the state of praxis up until the time that they are developed. To learn something about the

present shape and use of an artifact, a historical analysis of artifacts as well as of praxis is important (Engeström, 1987).

• Although collective, each activity is conducted through actions of individuals. Activity is what gives meaning to our actions, though actions have their own focus. Each action is implemented through a series of operations. Each operation is connected to the concrete physical or social conditions for conducting the action, and it is *triggered* by the specific conditions which are present at the time.

• Activities are not taking place in isolation but interwoven with other activities. Artifacts may be the instruments of a web of activities and in particular computer-based artifacts are often contributing to several activities or clusters of actions conducted by several users.

Human-Computer Interaction

To study human-computer interaction, some further concepts regarding human use of artifacts, and the mediating character of the artifact need to be elaborated. *Breakdowns* (Winograd & Flores 1986) related to the use process occur when work is interrupted by something - when the artifact stops mediating the work activity. In these situations the artifact as such or part of it becomes the object of our actions. *Focus shifts* generally occur throughout use. Breakdowns are one kind of focus shifts, and so are more deliberate changes of object, and articulation of the "otherwise unarticulated" in teaching situations. Through these concepts we are able to study the critical moments of the use of a certain piece of computer technology in a certain use activity; the critical moments when the computer application, or parts of it, looses its mediating character. Reminding ourselves that the categories of activity, action and operation can analytically be separated by the questions why, what and how, we may look at human-computer interaction as focusing on *how* a certain use of a computer application takes place. Thus, I will define the user interface as the parts of the computer software and hardware that support this how. The user interface is influencing which objects and subjects we can focus our actions on, while applying the artifact: the computer as a collection of buttons, the artifact as an object, other objects e.g. text documents, and so on. A good user interface will allow the user to focus on the objects or subjects that the user intends to work with. A bad user interface, on the other hand, will perhaps force the user to focus on other objects and subjects than the intended ones for major parts of the interaction. *Functionality* can be seen as something which only reveals itself in breakdowns and situations of reflection. In such situations it is possible to focus on *what* was done – presupposing a situation of use. Constituent parts of the user interface can be conditions both for operations directed towards the artifact, and for operations directed towards the real objects or subjects at different levels through the artifact.

An example of what can prevent us from shifting the focus from one subject/object to another, when intended, is that the parts of the application focusing on one subject or object are more or less segregated from other parts, focusing on other subjects or objects; they can be separate applications, running on the same computer. Also modes, and perhaps inconsistency of style of interaction, effect of commands, different icons, etc., can contribute to this segregation. Modeless interaction and integration of the parts of the user interface which in particular support actions and operations towards specific objects, can of course not, as such, prevent unintended shifts of object/subject.

We may characterize different aspects of the computer application based on the distinction between the different objects/subjects towards which the human being directs her actions and focus, and on the specific role and characteristics of these in the use activity. In (Bødker 1990) I distinguish between
• the *physical aspects*, the support for our operations towards the computer application as a physical object. We will meet this object in the total breakdown or before we get to know the application. The physical aspects are the conditions for the physical handling of the artifact. The human adopts to the forms and shapes of the artifact, and a mal-adaptation might prevent the forming of certain operations.
• the *handling aspect*s, the support for operations towards the computer application. A breakdown in these operations will make us focus on the artifact. The handling aspects are the conditions for the transparency of the artifact that allow the user to focus on the "real" objects and subjects of the activity. This type of operations can, for instance in breakdown situations, be conceptualized, whereby the user can be forced to conduct actions towards the artifact as an object.
• the *subject/object directed aspects* which constitute the conditions for operations directed towards objects or subjects that we deal with "in" the artifact or through the artifact. Different parts of the subject/object directed aspects relate to different subjects or objects, but it is also part of these aspects to support the shift between subjects/objects.

As an example we may look at two text editors for the Macintosh (MS WORD and MacWrite (Further details in Bødker, 1990)). The physical aspects of the two user interfaces are basically the same (the use of the screen and the mouse), although the two editors make different use of the keyboard. What we see on the screen, which in turn can help us recognize the program as a text editor, is in principle the same in the two cases. What makes the handling aspects are the "filters" put in between the user and the document: the direct representation of the document on the screen, and the scrolling mechanisms; the direct manipulation of the text (selection by dragging the mouse, cutting, pasting, etc.). When the user has used one of the programs for some time, there is no difference between

what she sees on the screen and the printed document, except for the distortion caused by the handling aspects. MacWrite, has more direct handling aspects than WORD. WORD on the other hand has more flexible handling aspects.

WORD seems to have better subject/object directed aspects than MacWrite, because it allows for more flexible formatting of the document – how we do things to the document. In none of the editors are there any direct support for operations towards other subjects or objects than the document.

We can make the following distinctions between different types of situations where the user is handling an object or subject through the computer based artifact:

1. The object is present only in the artifact (figure 1). An example of this is a spread sheet, which has no direct relation to objects outside the artifact (a print out of a spread sheet does not have the same capabilities as the spread sheet). The quality of the user interface must, for this type of artifact, deal with whether the user can distinguish between handling of the artifact and handling of the object in the artifact.
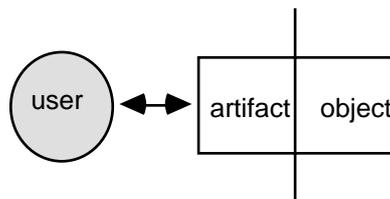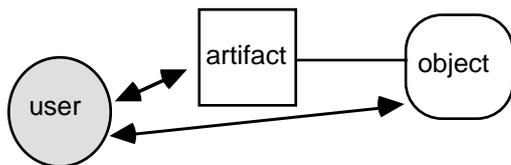

Figure 1. The object is present only in the artifact

2. The object exists as a physical object too, but is only present in the use activity as the representation in the computer application (figure 2). An example of this is a word processor: the object is a letter which is only present in the use activity as what can be seen and manipulated 'on the screen'. The quality of the user interface for this type of application must relate to how the user can couple the final object and the object 'on the screen' to each other. This is the type of situation where we can view parts of the user interface as a filter between the object as it is present in the use activity, and the real object.
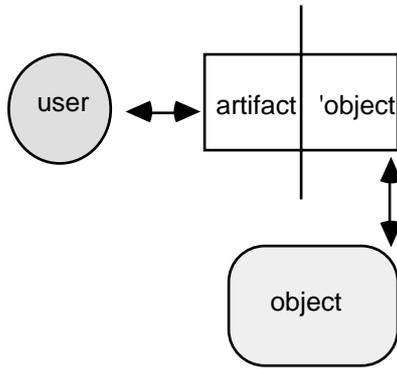
Figure 2. The object exists as a physical object, but is only present in the use activity as the representation in the computer application.

3. The object is present, physically, outside the artifact (figure 3). Examples of this are different kinds of control panels, where the object is handled through the artifact, but also physically accessible for inspection. For this type of user interface, too, the quality relates to the possibilities of coupling what is achieved through the artifact with what is happening with the real object.

Figure 3. The object is physically co-present outside the artifact.

Applications similar to the above types 2 and 3 exist for the support of the communicative side:
4. The other subject is not physically present in the use activity. An example of this is an e-mail system. Here, too, it must be possible for the user to make the coupling between how the other subject is experienced through the artifact, and the subject.
5. The subjects are physically present but communicate (partially) through the artifact. Examples of such applications can be found in the ideas of the Xerox Co-lab project (e.g. Stefik et al., 86) and many other meeting room facilities.

Summarizing, it is useful to see an artifact as supporting several interwoven activities, which deal with the same or connected objects: While conducting a specific activity, various focus shifts and breakdowns happen by which the object change. In some cases this may be viewed as a change of activity, in others the overall activity remains the same, but the purposeful actions change. There is a difference between actions directed towards (parts of) a computer-based artifact, which are indicators of

lacking transparency of the artifact, of insufficient handling and physical aspects of it, and actions directed towards objects in the actual realm of materials and products in the overall activity. Towards these, the user may direct actions or operations through the artifact, supported by the subject/object directed aspects of the artifact.

A case of situated analysis

In our current project[1] we work with the local branch of the Danish National Labour Inspection Service (NLIS). At the NLIS, a centralized system (VIRK) is applied to record the interaction of the NLIS with companies in the geographical area covered by the local branch. Visits to work sites as well as correspondence with companies are recorded, and various lists can be extracted, ranging from lists of a specific kind of companies within a geographical zone to lists of which recommendations and demands the NLIS has put on a specific company. Also lists of cases under investigation by a single NLIS inspector can be extracted.

The system was designed based on a company database shared with other authorities dealing with company inspection and counselling. It is a menu-based system running on terminals connected to a central database. The menu system is essentially hierarchical, where each mode allows for focus on one object of work. Focus shifts are supported mainly by traversing the hierarchical menu system or by short-cuts.

Historically, VIRK substituted a number of paper based lists, which were kept to maintain an overview of files concerning material about companies and inspections. These files are still used, only VIRK has made retrieval easier, and also some overview facilities for statistics have been added. Furthermore some of the paper-based lists are still maintained in situations where the support given by VIRK is too poor, e.g. in lists of various expiration dates, sorted according to expiration month is still kept, because VIRK offers little support for extracting such lists.

In our initial investigations we found that people used VIRK in many ways, and that some people actually asked for facilities which already existed in the system. Only very few people knew what VIRK really allows for. At the same time, it was our impression that VIRK could provide many of the functions that people asked for, only they didn't know how. As part of our project we sat out to find out how we could help the secretaries and inspectors at NLIS make better use of the system that they had already available.

The context of this specific piece of research is hours of interviewing the NLIS employees, of hanging around their offices and of partaking in seminars with them on other aspects of the project (see Bødker et al.,

---

[1] The project is conducted by Susanne Bødker, Ellen Christiansen, Pelle Ehn, Randi Markussen, Preben Mogensen, and Randy Trigg. For a description see (Bødker et al. 1991).

1991). All of these things have helped framing both problem with VIRK and the use of VIRK as such.

With the specific purpose of understanding VIRK and the HCI problems, three specific activities, all of which were videotaped, were in focus: A session with two secretaries where they discuss their daily activities, in particular with respect to documentation and information retrieval in VIRK, a session where a secretary is demonstration VIRK to the researchers, and where we ask questions to understand all parts of VIRK, and finally a session with a secretary who is also the "super user" of VIRK, doing the same. In total we have more that four hours of video-tape of the use of the system.

Starting out from the theoretical questions posed by activity theory, we have made an overall investigation of the setting in which VIRK is intended to be used, and its actual use setting. Furthermore a technique for mapping out the interaction between the user and VIRK was developed. As examples, some interesting situations were transcribed and mapped out accordingly.

Mapping it out

The initial analysis began by identifying the different levels of activities in which VIRK is applied: who are the users? what are the objects? to what extent are the objects present outside the computer application? which are the activities in which VIRK is used (why is a certain undertaking going on)? as to identify the role that VIRK plays in use (For details see Bødker, in preparation a and b). VIRK is the instrument of management of NLIS to make sure, that the people who contact NLIS get answered in due time. VIRK is used when distributing cases to inspectors and following up on their work. VIRK is used by the individual inspector and secretary to handle a certain case, and finally, VIRK is used by a secretary every time a document is registered in the system.

Generally, the objects that one can work on in or through the system has to do with recording the state of the overall activity. Descriptions and lists of documents, lists of cases, of deadlines, and various statistics, are the objects in VIRK. The contents of the cases, which are the objects dealt with by inspectors and secretaries when handling a case are almost absent in the system. (For an overview see fig. 4)

The analysis (see Bødker, in preparation a and b) helped see, that VIRK is designed as a planning and control system, which works rather well: management gets what it wants, and the data-entry works rather well to. VIRK is organized hierarchically, and it is hard to get from one function of the system to another. This is very much in line with the hierarchical way the organization as such is build up, and the division of work, when viewed only from the point of view of planning and control. There is a parallel between the intended division of work, and the separation of different parts of the system.

The needs and wishes of the secretaries and inspectors goes in the direction of integration and of tools for coordination of cases, for overviewing ones own cases, and for registering more informal and qualitative data about the cases. The facilities which are available in the system at present, are not very suitable for these purposes, and there seem to be no easy ways of extending the system to fulfil these needs. This distinction runs parallel with the historical account of NLIS given by Markussen (in press): NLIS has over the last twenty years moved from a workers' protection agency, where inspectors visit places of work rather randomly to enforce basic requirements to workers' safety, to an institution which applies a more holistic perspective on the work environment, where the inspectors take on also roles as therapists and teachers. At the same time the organization has moved from strict bureaucratic thinking to a client oriented perspective where quicker case handling, legible information, and uniformity of inspection is in focus. In many ways VIRK supports the bureaucratic aspects primarily.

This type of analysis helps explain why VIRK looks the way it does, and furthermore it can clearly be seen that there are two types of uses, which are running contrary to each other, these types of uses being fundamentally a matter of who is in control of the work situation: Where the systems perspective is a birds-eye perspective, viewing the human end-user primarily as somebody who provides accurate input to the system, in an interaction controlled by the system, the tools perspective focuses on the control of the human user over the tool as well as the material (This analysis expanded in Bødker, in preparation b).

VIRK can similarly be developed in both the direction of reporting and control, and in the direction of support for the work of the individual inspector as well as of groups.

For a more fine-grained analysis of use, the interaction was mapped out: The mapping consisted of listing in one dimension the focus during the session, and in the other the narrative of the situation. This narrative was supplemented with annotations of the users physical acting. The focus shifts appears as lines running from one "set of co-ordinates" to another. The focus can be categorized according to whether they have to do with the objects worked on (the subject/object directed aspects), or what should have been the handling of the artifact. Furthermore, the focus shifts can be categorized along the same lines: are they breakdowns caused by the work contents, the handling of the artifact of even by physical problems with the artifact. In the following example we can see what this means and where the analysis might take us (fig. 5).

The dashed arrow indicate that management is the acting user only by delegation. At no point in time is management physically touching VIRK.
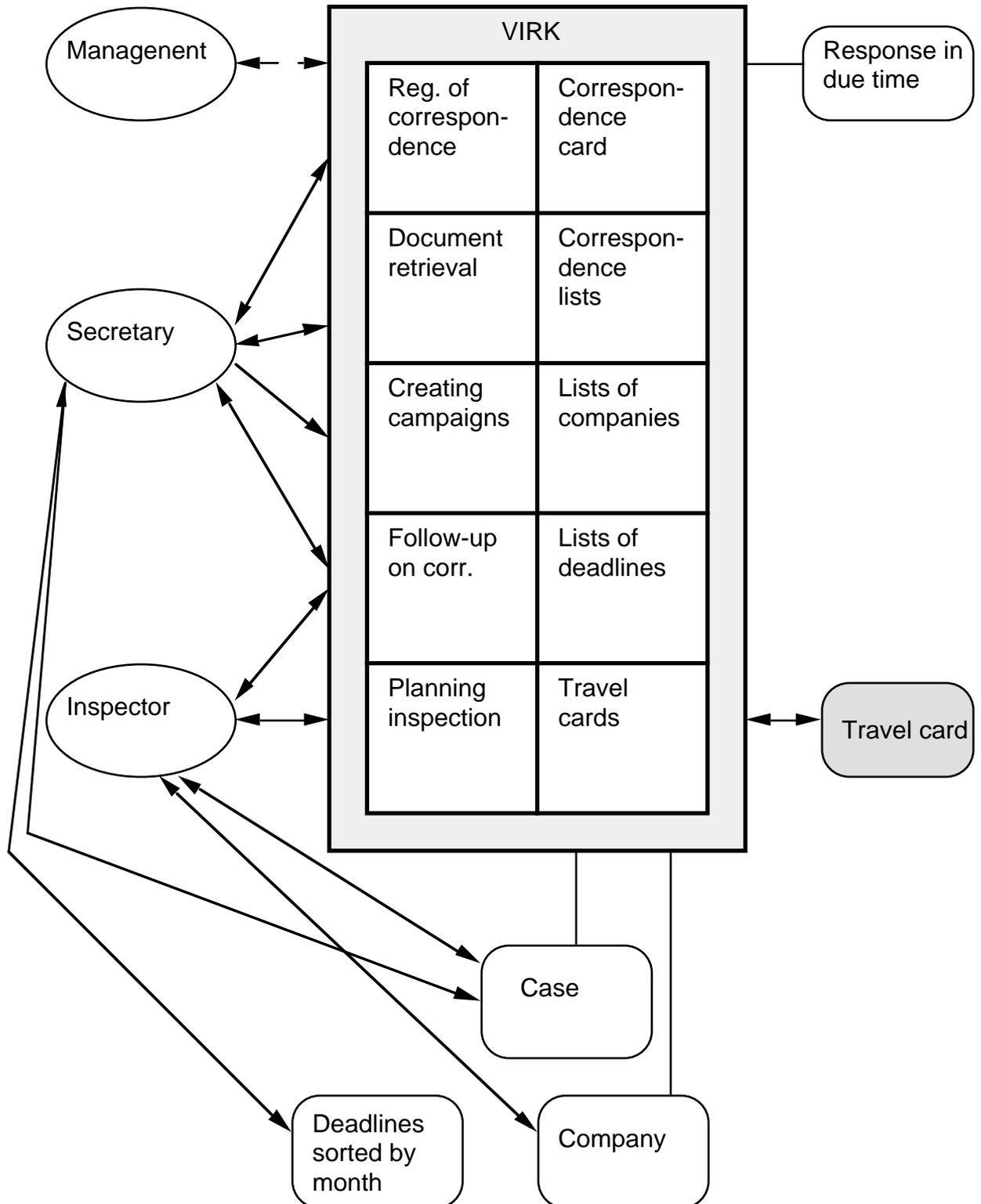

Figure 4. An overview of VIRK, its users, and the objects surrounding it.

Docu-ment	Field	Screen image	Report-generator	Report	Narrative
		(1)•			A: I am going to enter the (1)**next screen**. There it comes...where we can make this//
			HA		S: //there are some elevator programs there//
(3)•	SODA			(2)•	A:// (2)**this is something** we have made. [A report is already started]. (3)**This is what I want** sorted by registration number, this and that,
				(4)•	and then I know, let's see, then we want inspector (4)..**we extract accidents**...ways of sorting [A talks to herself] Then it is sorted based on numbers
		explanation, SODA			[Pause, waiting for result. Then long talk by A , cut out here]
		(5)•			...let me show you(5) **how we do that**. That is second search, every time I need it I must go back. You have to enter a (6)**sub screen**.
		(6)•			
to			HA	(7)•	S: then it is much more (7) **about how compare**
			HA (8)•		A: [typing search criteria] Look here, what the hell! here it is. No (8) **that is not it**, I am not...I don't have any
HA (10)•		(9)•			(9)**good material** about this. It needs (10)**experimenting**//
					S: //That is how it often goes
		(11)•			A: And then I need.. and (11)**then an x**.. then it should work. (12) **There it is**, that's how it looks.
		(12)•			

Figure 5. Generating a report

Legend: Bold face is used in legend to point at indications of a focus shift. the • indicate the object focused on, and the marking of the arches, the type of shifts in case.. A is the secretary, S the interviewer.

My interpretation of the above situation is the following: The focus is initially on the screen image, but as soon as A makes her choice, she is working on the report. The handling aspects of the artifact supports this focus. We have reason to assume that her focus was only on the screen image in the first place, because she was explaining VIRK to the researchers. She needs to specify which fields she wants. She has her focus on this, it is not done as an operation. This could again be because she is explaining, but the way she talks to herself indicates that she is

uncertain, i.e. she is conscious of the field because of a real breakdown. At the end of the session we see more breakdowns with respect to the handling aspects: she is rapidly moving between focuses on the report, the report generator, the field, in particular how criteria for this field is specified, and the written documentation and examples that she has available about the program. It is actually the handling of the search criteria that causes trouble: VIRK gives her no help in specifying neither syntax nor contents. At the end, she finds some sort of help in these examples, which allows her to focus back on the search criteria for the field, and over again at the report, and how the result looks on the screen. In the example we see, how focus shifts back and forth, where only the objects "field" and "report" has to do with contents or purpose of the activity as such. When trying to generate the report, the application is in no way transparent to A, which is seen by the numerous breakdowns with respect to the handling aspects, the many times where "screen image" and "report generator" comes into focus. In this case A is concentrating very much on doing the task, because it is so difficult, and the latter part of the example is a series of breakdowns with respect to the artifact. There is very little attention left for her to explain to us what she is actually doing.

Discussion
The above example is meant to illustrate that activity theory can be used as a basis for a rather detailed analysis of a situation of human-computer interaction, studying the role of the artifact-in-use in a real use situation. In Bødker, (1990, 1989, in preparation a and b) as well as here, I have argued strongly for the need to study artifacts in use. Whereas in my earlier work I have argued for the study of the technology in its actual work situations, (Bødker, in preparation a and b) have dealt with situations which deal with use, but are not real everyday use situations. The reasons for wanting to study use situations which are somehow artificial include to speed up the process and to get into critical situations. The major question is to what extent we can rely on these artificial situations to reflect problems of actual use?
From my examples (Bødker, in preparation a) I claim that it is possible to tell the difference between breakdowns of use and articulations of the otherwise unarticulated in relation to the role as teacher, that the user is taking on. This is because the framework allows us to tell the difference between different kinds of focus shifts and their relation to a real use activity. This means that there are shortcuts from the tedious studies of everyday work, but it does not mean that we can substitute real use situations with random demonstrations or testing by any random person walking in from the street. The user is in her normal environment, she has access to all her normal materials and facilities. Actually, it is only because the user knows her tasks as well as she does, and because the researchers know the contents and conditions of the work tasks as well as

they do that this works. This distinguishes this method from e.g. traditional thinking aloud experiments, which are often done in a more decontextualized setting, both physically and with respect to the praxis of the users and the pre-understanding of the analysts.

Reflecting on use of computer applications as a web of activities with different purposes, in which the computer application plays different roles rather than one helps focus on problems and contradictions in the interaction with the computer application.

Historical analyses of artifacts and work activities are useful in situating this analysis because it brings focus to the question of *why* use is organized the way it is, why the different roles of the artifact come into play, and in particular in understanding why some of the above contradictions occur. This *why* is reflected in what can be achieved by using the computer application and in how this is done; in the possible actions and operations undertaken by the user in each specific use situation.

My approach provides an alternative to traditional analysis of human-computer interaction which most often has no way of bridging to the specific use context of the artifact. What I have tried to show in this paper is that such a bridge can be developed when starting from human activity theory.

I have taken the first steps towards a more contextualized analysis of human-computer interaction. The human activity framework has provided a vehicle for bringing together an overall analysis of work and use of computer technology with a detailed analysis of human-computer interaction. It provides a way of understanding the relations between the overall socio-political conditions of work and the specific use of the specific computer application in a specific situation. What actually goes on may be analysed from many different levels of activity, where the actual actors are different, and maybe even different from those who conduct the specific actions. The mapping method is one way of coping with the complexity of real life work situations without reducing them to narrow, pre-defined laboratory experiments.

Outlook

Allen Newell once said that design is where the action is in the user interface. I completely agree that the real impacts of a better framework must, in the end, be demonstrated in better design of computer applications (For a further discussion, see Bødker, 1990). How can the framework be put into action in designing better user interfaces? The challenge for the design of an artifact is to build on existing use praxis. Thus, it must be a deliberate effort in design not only to build new artifacts but to help the future users reshape their praxis.

Use, as a process of learning, is a prerequisite to design. Through use, new needs arise, either as a result of changing conditions of work or as a recognition of problems with the present artifacts through breakdowns. In

design the future users must be able to assess the artifact-to-be in use. Thus, the designers must be able to take part in design and help interpret breakdowns of use. I have suggested (See Bødker, 1990) that prototyping, using computers or mock-ups, is a useful way of working with this issue. Such an action-oriented approach to human-computer interaction furthermore fits in very nicely with action-oriented approaches to work research as suggested by e.g. Engeström (1987).

One may well argue that my approach to activity theory is a very pragmatic one. Rather than involving myself in a critical discussion of the differences between (different schools of) activity theory and my other sources of inspiration (Winograd & Flores, ethnomethodology), I have discussed some ideas from activity theory that can inform our praxis as researchers in human-computer interaction. Strongly inspired by Mike Cole, I would like to stress that I see approaches such as this as important ways of proceeding activity theory in the years to come.

References

Bannon, L. and Bødker, S. (1991). Beyond the Interface: Encountering Artifacts in Use. In J. M. Carroll (Ed.) *Designing Interaction: Psychology at the Human-Computer Interface*, New York, Cambridge University Press, pp. 227-253.

Bødker, S. (1989). A Human Activity Approach to User Interfaces, in *Human Computer Interaction*, T. Moran, (Ed.), Vol. 4, No. 3.

Bødker, S. (1990). *Through the Interface – a Human Activity Approach to User Interface Design*, Hillsdale, NJ, Lawrence Erlbaum Associates.

Bødker, S. (in preparation a). Understanding computer applications in use - a human activity analysis. In P. Bøgh Andersen, B. Holmquist H. Klein & R. Posner. *The semiotics of the workplace*, in preparation.

Bødker, S. (in preparation b). *Historical analysis and conflicting perspectives - contextualizing HCI*. East West HCI 93.

Bødker, S., Christiansen, E., Ehn, P., Markussen, R., Mogensen, P. and Trigg, R. H. (1991). Computers in Context. Report from the AT-project in Progress. *Report of the 1991 NES-SAM conference*, Ebeltoft, Denmark.

Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine – the power of human intuition and expertise in the era of the computer*. Glasgow: Basil Blackwell.

Engeström, Y. (1987). *Learning by expanding*. Helsinki, Orienta-Konsultit.

Markussen, R. (in press). A historical perspective on work practices and technology. In P. Bøgh Andersen, B.Holmqvist, J. F. Jensen (Eds.): The Computer as a Medium, Cambridge University Press.

Stefik, M., Bobrow, D. G., Lanning, S., & Tatar, D. (1986). WYSIWIS revisited. Early experiences with multi-user interfaces. *Proceedings from CSCW '86. Conference on Computer-Supported Cooperative Work, December 3-5 1986 Austin Texas* (pp. 276-90).

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge: Cambridge University Press.

Winograd, T., & Flores C. F. (1986). *Understanding computers and cognition: A new foundation for design*. Norwood, NJ: Ablex.