3

Computer-Mediated Activity: Functional Organs in Social and Developmental Contexts

Victor Kaptelinin

The field of human-computer interaction (HCI) presents an enormous theoretical challenge to researchers trying to establish it as an integrated field of studies. To become such a field, HCI should be based on a conceptual scheme powerful enough to incorporate both human beings and computer technology within a coherent theoretical framework. One possible solution is based on the cognitive approach, according to which both human beings and computers can be considered as information processing units.[1] If the basic mechanisms underlying human cognition and those underlying the functioning of computer systems are essentially the same, it is possible to use the same concepts and methods to analyze both entities and eventually to build a general theory that explains the functioning of higher-level systems composed of both human beings and computers.

Another broad approach, which is becoming more popular, is based on a radically different assumption. It assumes that what is needed to make HCI a conceptually integrated field is a theory that describes and explains the larger context of human interaction with computers. This second approach employs another feature human beings and computers have in common: both are involved in real-life activities of computer use. If we can provide an account of the general context of computer use and identify the place of human beings and computers within this overarching scheme, we can understand interaction between them without assuming that they are basically the same kinds of entities.

There are several versions of the contextual approach to HCI (see Nardi, chapter 4, this volume), which are almost unrelated to each other. What they have in common is their opposition to the currently dominant cognitive approach and related methodology and the more or less explicitly formulated idea that both human beings and computers develop in the process of cultural history and can be understood only within a social context.

Activity theory is one of the concrete versions of the contextual approach. The theory is becoming more talked about in the field of HCI, but is still ``opaque'' for most researchers (Brusilovsky, Burmistrov, and Kaptelinin 1993). This chapter discusses the potential advantages and limitations of activity theory as a conceptual framework for HCI. It is not intended to present the basic ideas and principles of activity theory in relation to HCI (for such an introduction, see Bødker 1989, 1991; Kaptelinin 1992, this volume; Kuutti 1992, this volume; Zinchenko 1992, this volume) but instead attempts to put the theory into the context of the problems that researchers in the field are currently encountering.

EXPANSION OF HUMAN-COMPUTER INTERACTION

One of the most important claims of activity theory is that the nature of any artifact can be understood only within the context of human activity—by identifying the ways people use this artifact, the needs it serves, and the history of its development. Activity theory itself is a special kind of artifact. That is why it is important to understand the motivation behind the actual and potential use of this theory as a conceptual tool in the field of HCI. It can be useful in developing realistic expectations about the scope and potential outcomes of the theory.

Considering activity theory as a special kind of tool implies that accepting this perspective does not exclude other approaches and does not reject the usefulness of other conceptual schemes (because no tool, no matter how powerful it is, can serve all needs and help to solve all problems). In particular, activity theory does not reject the value of cognitive studies. However, the general conceptual position of activity theory is radically different from that of the cognitive approach. Specifically, activity theory does not allow for an equal status of human beings and computer technology in a theoretical framework of HCI. The relations between agents and tools cannot be symmetrical, and this fundamental fact should be taken into account in developing a theory of HCI.

The history of studies in HCI clearly demonstrates the tendency of ever-extending units of analysis (Grudin 1990). While the early attempts to understand the factors influencing human-computer interaction concentrated on low-level input-output processes, the current focus is on long-term events and large-scale aspects of HCI, such as the software development life cycle, computer-supported cooperative work (CSCW), and the implementation of information technologies at the organizational level. The reasons behind this tendency are not only theoretical; they also include practical considerations. It has turned out that the quality of the user interface and of interactive systems in general depends on factors that are in no way limited to the sensorimotor level of interaction (although the latter is still important).

The challenge of increasing units of analysis is faced by any theory of HCI. The solution offered by both the cognitive approach and activity theory is to consider human interaction with computers as a multilevel hierarchical structure. According to both the cognitive approach and activity theory, the tendency to focus on the higher-level HCI events is related to taking into consideration the higher levels of the hierarchical structure of computer use. However, as we will see, the meaning of hierarchical organization is different within these two approaches. According to the cognitive approach, the major theoretical task is to develop a conceptual scheme that can give a coordinated description of multilevel information processing in both human beings and computers. According to activity theory, the hierarchical organization of human-computer interaction is determined by its embeddedness into the hierarchical structure of human activity that mediates the user's interaction with reality.

The current expansion of the subject matter of HCI can be described as a three-dimensional ``explosion'' of the traditional paradigm, with ``levels of interaction'' being only one of these dimensions. Other dimensions, difficult to account for within the cognitive approach, are set out in figure 3.1.

First, there is a shift of focus from interaction between the user and the computer to a larger context of interaction of human beings with their environment, that is, transcending the user interface to reality beyond the ``human-computer system.'' Computer tools are used by people to reach meaningful goals that usually exist beyond the situation of human-computer interaction, and, moreover, often serve as intermediate steps to higher-level goals that can be even more remotely related to the situation of computer use. This aspect was described in some ethnographic studies of the use of technology (Suchman 1993). It was shown that interactions with information technology are embedded into logically structured sequences of interaction with other objects and with people.

The next dimension is that of development. The components of human-computer interaction are not static. The user begins as a novice and often ends up an expert; assimilation of new artifacts can solve old problems, but changing the nature of the tasks performed by the user, it creates new problems that require still new artifacts to be used (Carroll, Kellogg, and Rosson 1991). An understanding of a particular case of computer use includes an analysis of its history and its potential developmental transformations.

Finally, there is the individual/social dimension. The current meaning of the word *user* now includes not only individuals but also groups and organizations.

The discovery of the rich, multifaceted, and multidimensional reality of human-computer interaction is probably one of the most salient features of the current situation in the HCI community (Bowers and Rodden 1993; Kuutti and Bannon 1993; Russell et al. 1993). The powerful expansion of the object of study opens important new horizons—and at the same time, creates a feeling of confusion. The field of HCI seems to be a collection of loosely related subfields; familiar concepts suddenly turn out not to be so simple anymore. (This feeling of confusion is evident in, for example, Bowers and Rodden 1993 and Monk et al. 1993.)

This situation in the HCI community constitutes the context that can explain the growing interest in activity theory. The new reality of human-computer interaction requires new theoretical tools to help overcome the ``explosion'' of the subject matter of HCI, to coordinate the efforts of the increasing number of researchers working in this interdisciplinary field, and to find a way to make the outcomes of the studies more relevant to practice. It is natural and logical to try various theoretical approaches that can potentially provide a consistent picture of the field, and activity theory seems to be one of the most promising candidate approaches. It is true that activity theory is not a ready-made universal solution to all the problems of human-computer interaction, but it is also true that the general vector of the current development in the field of HCI (Bannon 1991) is directly related to the very essence of this theoretical approach.

COMPUTER TOOLS AND FUNCTIONAL ORGANS

One of the most salient features that distinguishes activity theory from the cognitive approach is that activity theory considers computers as a special kind of tool mediating human interaction with the world. Meaningful, goal-directed activities constitute the context for both mental processes and external actions. Human beings usually use computers not because they want to interact with them but because they want to reach their goals beyond the situation of the ``dialogue'' with the computer. As formulated by Bødker (1991), users are acting ``through the interface.'' Therefore, the subject matter of HCI should not be a closed system of ``user-computer'' but should include the meaningful context of the user's goals, environment, available tools, and interactions with other people.

The need to expand the object of analysis is dictated not only by general theoretical interests but also by specific design considerations. It is not possible to create a high-quality system while relying solely on abstract universal guidelines (such as logical consistency) and ignoring the larger context of human activity. As Grudin (1989) showed, designers sometimes deliberately violate the principle of logical consistency to make systems more usable. They do it to make systems more consistent with the general structure of the user's activities, and the logic of consistency in this sense can be different from the logic of internal consistency.

The tool mediation perspective has important theoretical implications for HCI. First, it questions the very name of the discipline. The emphasis on the interaction implied by this name seems to be a little misleading. The chronological sequence of terms used to cover the problems of HCI demonstrates the more and more sound accent on human activities rather than on computer systems (``human factors,'' ``computer and human interaction,'' ``human-computer interaction,'' ``computer-supported cooperative work''). It would not be surprising to see a new term for the discipline, stressing the tool nature of computers (for example, ``computer-mediated activity''), and even if the discipline retains its current name, it will probably be used just as a label that has little to do with its actual content, just as ``computers'' are no longer associated simply with calculating devices.

Second, if we accept the tool mediation perspective, we have to deal with two interfaces instead of one user interface, with two borders, separating (1) the user from the computer and (2) the user *and* the computer from the outside world (Bødker 1991; Grudin 1993; Norman 1991). This duality raises a problem that is almost identical to Bateson's (1972) blind man's stick dilemma: where is the boundary between the individual who uses a tool and the external world? Does it coincide with the individual-tool boundary or with the tool-world one?

The activity theory answer to this question is based on the concept of ``functional organs'' (Leont'ev 1981). Functional organs are functionally integrated, goal-oriented configurations of internal and external resources. External tools support and complement natural human abilities in building up a more efficient system that can lead to higher accomplishments. For example, scissors elevate the human hand to an efficient cutting organ, eyeglasses improve human vision, and notebooks enhance memory. The external tools integrated into functional organs are experienced as a property of the individual, while the same things not integrated into the structure of a functional organ (for example, during the early phases of learning how to use the tool) are conceived of as belonging to the outer world.

Computer tools share the common attribute of all tools: they are integrated into functional organs. From the point of view of activity theory, the nature of these functional organs is of special interest to HCI studies. Perhaps the central problem of HCI can be defined as that of optimal integration of computer tools into the structure of human activity (Kaptelinin 1992a). What are the needs that require the development of a new functional organ? What is the range of goals that are intended to be reached with the new tool? What is the structure of human activity before the assimilation of the tool, and what is the previous experience of computer users with the uncomputerized equivalent of this activity? All of these questions are no less relevant to the ``transparency'' of the user interface than an optimal width and depth of the menu system (Shneiderman 1987).

Certainly there are several kinds of functional organs based on the use of computer tools, because computer tools do not have one fixed function. One of the most important functions can be defined as an extension of the internal plane of actions (IPA; figure 3.2). The IPA is a concept developed within activity theory that refers to the human ability to perform manipulations with an internal representation of external objects before starting actions with these objects in reality. It is similar to the cognitive concepts of working memory and mental models, but it refers not to specific mental models but to the general ability to create and transform them.

The IPA appears at a certain stage of child development (the most critical period corresponds to the early school years, according to Ponomarev 1975) and constitutes a new kind of interaction between internal activities and external ones. Human activities include external and internal components at every developmental stage. Initially, however, the function of internal components is limited to the control of external activity; that is, the only way to get feedback to an action is to perform the action in reality. Over the process of internalization—the transformation of external activities into internal ones—the child acquires an ability to perform some actions ``in mind'' and in this way avoids costly mistakes and becomes free from the immediate situation. A system of mental structures and abilities that makes it possible to perform actions ``in mind''—the IPA—is the result of this development. The same control mechanisms that regulate external behavior can provide feedback on actions performed in the IPA (actually, these control mechanisms also develop, but this is a separate problem not directly related to this chapter). Oversimplifying, we can differentiate between two subsystems of human cognition: central cognitive structures and the IPA (see figure 3.2).

The potential of computer systems to create easily controllable models of target objects and to give the user the opportunity to evaluate them and to manipulate them explains why such applications as spreadsheets, word processors, and graphics editors became so popular.

Computers are not the only kind of tools used as an extension of the IPA, and this function is not the only function of computers (they are also used, for example, as communication tools), but an understanding of the mechanisms underlying the use of computer tools as extensions of the IPA is directly relevant to the development of useful and usable systems. Some recent systems have excellent representational facilities supporting experimentation with individual objects (such as the Print Preview window in Microsoft Word), but there is still much to be done to make computer tools more efficient extensions of the IPA. In particular, existing applications have rather limited potential to support comparison of multiple objects or to give a full picture of large objects. Of course, the screen size imposes severe limitations, but this problem can be partly solved by intelligent use of zooming and simulation of information integration during saccadic eye movements (after all, the focal vision of human beings is also very narrow). Another opportunity to make computer tools more adequate to their role as the IPA extension would be to support a backtracking search strategy, to provide the user with a representation of the search process history, to help the user in evaluating the results, to support coordination of the given task with other kinds of activities performed by the individual at the same time (for example, to give an estimation of the time needed to print out the specific document on the specific printer). In short, understanding the requirements to computer-based functional organs, as well as the mechanisms underlying the integration of computer tools into functional organs, can stimulate insights into the practical problems of HCI, problems related to the design of new kinds of interactive systems.

Finally, the tool mediation perspective in HCI brings to the field the issue of culture. Tool mediation is a way of transmitting of cultural knowledge. Tools and culturally developed ways of using tools shape the external activity of individuals and through the process of internalization influence the nature of mental processes (internal activity). The role of tools is not limited to transmission of operational aspects of human interaction with the world. As Latour (1993) emphasized, tools also shape the goals of the people who use the tools. There are implicit goals that usually are ``built into'' the tools by their developers. The goals achieved by people equipped with a tool are often influenced by the ``tool's goal,'' and the final results differ from both goals, being a compromise between them. (According to Latour, the person who has a gun can be influenced by the implicit ``goals'' of the gun even if the gun is never used.) The same is applicable to computers and software. The values and goals intended by their developers can influence users who may not even be aware of these influences. This is obvious in the case of some computer games but might be true with respect to other kinds of applications, too; for example, the style of communication via e-mail can be influenced by the nature of this medium, or a database format can influence the way people differentiate between important and less important facts.

The tool perspective in HCI calls for a revision of many traditional concepts (e.g., the concept of ``interface'') and raises many problems, including the mechanisms underlying the integration of computer tools into the structure of human activity (functional organs) and the coordination of general cultural perspectives of the people involved in the development and use of computer tools.

DEVELOPMENT OF COMPUTER-MEDIATED ACTIVITY

Assimilation of computer tools, by either individual users or organizations, is a continuing process rather than a single act. The need for the constant change stems from several sources. The first is related to technological progress. New generations of hardware and software change standards and requirements for computer tools. To keep up with technological development, users have to adopt newer systems. Another source is related to the developing needs of users. The use of a particular tool changes the structure of activity and can result in new goals to be satisfied. This phenomenon, described by Carroll, Kellogg, and Rosson (1991) as the ``task-artifact cycle," can be illustrated with numerous examples. Sometimes it is very difficult to predict the line of development, and in some cases it can hardly be characterized as progress toward more powerful tools. Several years ago the interface of an e-mail system I was using suddenly changed. The machine became so popular (it was one of the first e-mail servers in the former Soviet Union) that the system administrators had to provide most users with a very primitive interface that allowed them to save messages only as a file on the user's floppy disk and to send messages stored in a prespecified format as files on the floppy. This design was intended to limit the time the users spent working at this particular computer by preventing them from the use of advanced facilities of the system.

The changing requirements for computer tools raise a special problem of making it possible to meet these changing requirements efficiently. For this reason it is practical to be able to predict the changes or at least the general tendencies. Of course, it is important to make systems flexible, so that users can adapt them to their needs, but it would be a mistake to leave the problem of adaptation to users only. Many popular systems provide users with an opportunity to change the interface according to their wishes. However, a number of studies have shown that most users do not utilize these facilities at all, rather, they need help in implementing the necessary changes to the system.

The cognitive approach does not provide any substantial basis for solving this problem. Activity theory, which distinguishes between various levels of determination of the agent's (individual's or group's) behavior, can give some hints. The most general idea is that understanding the status of a process within the conceptual structure of activity can help to anticipate the direction of potential changes, as well as related costs and benefits. If a change is limited to the level of operations, the problems associated are technical ones (financial resources needed, time necessary for reautomatization of routines, etc.). If, however, it turns out that some goals are no longer meaningful, a careful analysis of motives impelling the individual and/or organizational activity, as well as alternative ways to reach this motive under the current circumstances, should be conducted. The most difficult problems arise if the changes reach the level of the whole activity. The changes of activity structure caused by the use of computer tools usually take place at different levels simultaneously.

It would be an exaggeration to say that the problem of development has not been studied within the cognitive approach. There are numerous studies of differences between novice and expert users (Allwood 1989) and many cognitive models explaining the mechanisms underlying skill acquisition (Anderson 1983). From my point of view, however, cognitive models cannot efficiently deal with qualitative changes of cognitive skill with practice.

Recently I (Kaptelinin 1993) investigated the phenomenon of poor recall of command names in pull-down menus by experienced system users, as discovered by Mayes et al. (1988). Subjects first practiced with a very simple menu-based system and were then exposed to two transformations of the initial interface: (1) all the command names were substituted with strings of dots but their order was the same, and (2) the items were scrambled within each menu during every new task. I found that after the initial loss of efficiency, subjects could quickly restore speed and accuracy level in the case of the ``dotted" menus but not in the case of the scrambled ones. So while during the early phases of learning the menu selection was based on the command names, during a relatively short period of time of initial practice (about one and a half hours) the subjects seemed to switch from verbal clues to menu selection based on spatial locations of the items.

These data illustrate the complex nature of developmental skill transformations, which cannot be completely explained in terms of ``chunking" or ``knowledge compilation." They are also relevant to the concept of ``affordances" as it was introduced to the field of HCI by Norman (1988). There is no doubt that affordances are very helpful if the situation is not familiar to the user (as during the exploratory learning of a system) or the goal is unambiguously determined by the situation, which does not allow for a wide range of possible actions (e.g., the goal is to open the door I want to go through), or both. However, the benefits of affordances beyond these limited conditions are not so obvious. The notion of affordances implies that the user is matching his or her goal against the set of opportunities offered by the environment, that he or she directly sees what can be done to reach the goal. In other words, the notion of affordances implies that the objects have some universal operational meanings (what can be done with these objects) that is directly communicated to the users. From the point of view of activity theory, human beings actively create the meaning of the objects in the process of interaction with the environment. This idea can be illustrated with the study of menu selection: the users ended up relying on spatial locations, which cannot be considered as affordances at the early phase of interaction with the system.

Besides, in most real-life situations there is no one-to-one correspondence between goals and ways to accomplish the goals. There are often many ways to achieve a goal. The use of affordances implies that the initiative, in a sense, is taken by the external situation. According to activity theory, the elementary components of activity—operations—are not just triggered by conditions, they are determined by the general structure of the action they are incorporated into. People learn to control their immediate impulses, and an important aspect of social norm acquisition by children seems to be learning to ignore some affordances.

GROUP AND ORGANIZATIONAL COMPUTER USE

Since the cultural-historical tradition represented by activity theory emphasizes the social nature of human beings and their activities, it appears natural to expect the most tangible benefits from activity theory in studies of social aspects of computer use, for example, in the field of CSCW. From my point of view, it would be too optimistic to think of activity theory as an approach that can provide ready-made answers to the problems related to group and organizational computer use. However, it appears that basic principles of activity theory can be elaborated on and operationalized to make the theory a useful tool for studying supraindividual levels of information technologies use. This section advances some arguments supporting this point of view.

Activity theory has been developed as a psychological approach, and it almost exclusively deals with individual human beings. Undoubtedly, social context plays an important role in activity theory, but it is mainly used to explain how individuals are influenced by social factors, not to give an account of activities of social units. However, the general notion of a developing active agent interacting with an environment in a social context is applicable not only to individuals but to groups and organizations as well.

Several attempts have sought to expand the concept of activity to supraindividual phenomena. One of them was made by Engeström (1987), who proposed a scheme of activity different from that by Leont'ev (1978, 1981); it contains three interacting entities—the individual, the object, and the community— instead of the two components—the individual and the object—in Leont'ev's original scheme. Engeström's version of activity theory has been adopted in some recent studies in the fields of HCI and CSCW (see Kuutti, this volume).

Another attempt to extend activity theory beyond the individual level was made by Soviet psychologists in the seventies and eighties (Petrovsky and Petrovsky 1983). The concept of collective subject (*kollektivnyj subjekt dejatelnosti*) was introduced to account for the processes of communication between individuals. According to this approach, communicative processes can be conceived as interactions between structural components of the collective subject, the interactions that are subordinated to the primary kind of interaction—the activity relating the subject to reality. The explanatory potential of the concept of collective subject is somewhat limited since there definitely are communication phenomena that do not fall into the category of interactions between members of a team pursuing a common goal. On the other hand, there no less definitely are phenomena of communication and cooperation that do fall into this category, and within its scope, the concept of collective subject can be useful.

Computer-supported activity of a group or organization can be analyzed along the general lines of activity theory: finding the motive, goals, and conditions of the activity; identifying structural components of the subject's interaction with reality (individual activities, actions, and operations) as well as tools mediating the activity; and tracing developmental changes of the activity. There are, however, some activity theory concepts whose meaning is not clear in the context of collective activity; it is not even clear whether they can be applied in this context.

One difference between individual and collective activities is rather obvious: a structural component of a collective subject can be a subject, too, with his or her own motives and goals. This difference is essential, yet it is still possible to address the problem of group activity decomposition from the standpoint of activity theory. According to Leont'ev (1978), actions are usually polymotivated; two or more activities can temporarily merge, motivating the same action, if the goal of one action is a prerequisite for reaching the motives of all of the activities simultaneously. This principle can also be applied to integrating individuals or groups within the structure of a higher-level collective subject; it is not necessary that all component subjects share the motive of the system they are incorporated into, but the goals of the subjects should permit polymotivation, that is, should satisfy motives of both the component subject and the system. A specific case of the polymotivation principle is ``Grudin's law," as formulated by Norman (1993): ``When those who benefit are not those who do the work, then the technology is likely to fail or, at least, be subverted."

The notion of computer tools as extensions of IPA can be applied to collective agents, too. In this case, the components of group activity can be presented as shown in figure 3.3. Shared virtual workspace is an extension of the group IPA, and the pattern of various links relating the components of the model represent the way computer-mediated communication, noncomputer communication, interpersonal communication mediated by references to a virtual common object, and individual human-computer interactions are integrated within the collective computer-mediated activity. Figure 3.3 is not intended to represent all aspects of the complex reality of computer-mediated group work, but it can help to relate individual and collective computer use and by this means contribute to bridging the gap between traditional HCI and CSCW (see, e.g., van der Veer 1994).

CONCLUSIONS

The difference between activity theory and the cognitive approach has been analyzed in relation to such aspects of human-computer interaction as affordances of the user interface and computer-mediated group activity. This concluding section will present two more examples of differences between the approaches and then, will focus on limitations of activity theory as a conceptual framework of human-computer interaction.

Both activity theory and the cognitive approach view human-computer interaction as a hierarchically organized process. However, there is an important difference between the cognitive approach to levels and the activity theory approach. The highest level the cognitive models are dealing with—the task level, according to Moran (1981)—is usually barely mentioned. The details of the dialogue structure can be ideally deduced from a higher-level specification of a system, but it is very difficult (if not impossible) to give a cognitive explanation of how this higher-level specification is being created by designers.

This point can also be illustrated with the GOMS model (Card, Moran, and Newell 1983), which is intended to be used in evaluating user interface designs. There are some striking similarities between activity theory and GOMS.[2] Both approaches are intended to give a description of goal-directed behavior. The use of the term ``goal" in activity theory corresponds to the use of ``goal" in GOMS. Other correspondences are between ``operators" (GOMS) and ``operations" (activity theory), ``methods" (GOMS) and ``actions" (activity theory). However, a closer look at these approaches reveals fundamental differences between them. First, according to activity theory, the relations between actions and operations are *dynamic*: actions become operations through automatization, and operations can transform into actions in cases of breakdown. The GOMS model, however, deliberately avoids considering nonroutine processes and does not even intend to describe the dynamics of the interaction structure. Second, activity theory describes both the subjective and the objective sides of interaction, and goals and conditions are understood as aspects of reality that direct and constrain the actions and operations of the individual. The GOMS model, on the other hand, is intended to represent individual behavior abstracted, as much as possible, from the environment. Third, and most important, activity theory puts goals and actions into the context of activities, while GOMS does not deal at all with the origins of goals (the model is just not intended for this purpose). Before applying the GOMS approach, one defines the goals. So, in the area of user interface evaluation, an application of cognitive models raises the same problem as in the area of the user interface design: how to identify the basic aspects of the system to be created, how to capture the everyday experience (actual and potential) into a description that can be used as a starting point for system development.

One potential solution would be to combine radically different approaches in proceeding through two major phases of creating system specifications. A contextual approach (for example, ethnography) can be applied at the first phase to analyze the everyday environment of the potential users, to understand their needs, and to take into account all possible factors that can influence the quality of the target system. The outcomes of this phase, a general specification of the system, can be passed to the second phase, where a version of the cognitive approach can be employed to specify the details of the system (this specification, in turn, can be used to start the whole process of iterative design).

Unfortunately, this ideal scheme has serious drawbacks. First, the mediating specification should be understandable for people working in different paradigms. It is not clear if this is possible at all (Monk et al. 1993). Second, the process of stepwise refinement of the system specification is not a strictly top-down one; important constraints are sometimes discovered at lower levels and make it necessary to come back to higher-level decisions and correct them so that it is possible to meet these constraints. Therefore, the coordination of various levels of specification could require continuous back-and-forth coordination of different perspectives and is a potential source of numerous problems. That is why it is desirable to rely on a single homogeneous conceptual scheme powerful enough to cover various levels of human-computer interactions. Activity theory integrates multilevel perspectives on human activities within a single conceptual framework, and this is probably the main reason it has attracted the attention of many researchers.

The difference between the cognitive approach and activity theory can also be illustrated by the interpretation of artifacts within the cognitive paradigm. Probably the most advanced attempt to introduce the idea of mediation into the cognitive approach was made by Norman. In his paper ``Cognitive Artifacts," Norman (1991) points out the importance of analysis of a special kind of artifact that he defines as follows: ``A cognitive artifact is an artifact designed to maintain, display, or operate upon information in order to serve representational function." The introduction of this concept leads Norman to differentiate between the personal view and the system view of human-computer interaction. These views correspond, respectively, to the border between the user and ``the computer and the world" (the personal view), and the border between ``the user and the computer" and the world (the system view). Norman concludes that the notion of empowering people with computer tools is true only from the system view, while from the personal view, the use of computers means just changing the nature of the task (which can become easier and require fewer capacities from the user).

These conclusions can be compared to the activity theory interpretation of the same phenomena. First, activity theory does not differentiate between the ``personal view " and the ``system view." Both views are considered personal; the functional organs employed in the two cases are, of course, different, but both are functional organs of the individual, and hence personal. Second, and more important, activity theory states that tools not only change the task but often empower the individual even if the external tool is no longer used. Numerous experiments have shown that activities mediated by symbolic tools often undergo three developmental stages: (1) the initial phase, when performance is the same with and without a tool because the tool is not mastered well enough to provide any benefits, (2) the intermediate stage, when aided performance is superior to unaided performance, and (3) a final stage, when performance is the same with and without the tool but now because the tool-mediated activity is internalized and the external tool (such as a checklist or a visualization of complex data) is no longer needed. Even if the tool does not seem symbolic at all, its external use can substantially influence the nature of internal activities. For example, a novice pool player has to hit a ball with the cue to see the results; an expert may not need the physical cue (with which to make the shot) to know the results of a planned shot.

The difference between Norman's view and that of activity theory illustrates that the principles underlying activity theory are closely interrelated. The concept of cognitive artifacts that Norman introduced is almost identical to the concept of psychological tools developed in the cultural-historical tradition (see, e.g., Engeström 1987), but the use of the concept out of the context of the other principles leads to conclusions opposite those of activity theory.

In this chapter, I have emphasized the potential advantages of activity theory rather than its disadvantages (since my point is to substantiate further attempts to apply activity theory to the problems of HCI, as well as to outline the most promising lines of such an application). But this emphasis does not imply that activity theory can provide solutions to all problems. Activity theory has very serious limitations too. Let us consider some of them.

First, activity theory was mainly developed as a psychological theory of individual activity. This is an important limitation, because the current meaning of the term ``the user" includes not only individuals but also groups and organizations. Many researchers agree that activity theory can be applied to supraindividual units, such as groups and organizations; however, the specific conceptual system necessary for analysis of social systems is still under discussion. In the former Soviet Union, the opportunity to study social phenomena was limited for political reasons. Probably the only relevant idea developed by the Soviet proponents of activity theory was the notion of collective subject. This concept is less elaborated compared to the aspects of activity theory related to individual subjects, and it remains to be clarified to what extent the conceptual apparatus of activity theory is applicable to collective subjects. Important developments toward the extension of activity theory to the level of social processes have been made by Western researchers (Cole 1984; Engeström 1987; Raeithel 1992), but the problem is not solved yet.

Second, compared to the cultural-historical approach developed by Vygotsky (1978), activity theory adopted a narrower point of view of culture. Activity theory was oriented to practical needs of society, was greatly influenced by the example of natural science, and always tended to interpret reality in formal schemes (see Zinchenko 1992; this volume). While culture, values, motivation, emotions, human personality, and personal meaning are embraced by the conceptual system of activity theory, the theory does not aim at giving a comprehensive description of all these phenomena. It captures only some of their aspects: those related to rational understanding of human interaction with the world. This feature of activity theory can be considered a benefit because it is similar to the way many system developers think, but it might also be viewed as a disadvantage, because activity theory cannot completely substitute for an anthropology that defines and understands culture.

Third, the tool mediation perspective, which is considered the most important advantage of activity theory, can also impose some limitations on its potential application. In virtual realities, for example, the border between a tool and reality is rather unclear; information technology can provide the user not only with representations of objects of reality but also with a sort of reality as such, which does not obviously represent anything else and is intended to be just one more environment with which the individual interacts. Virtual realities present a problem to activity theory that probably cannot be solved without enriching activity theory's basic principles with new ideas from the cultural-historical tradition or other related approaches.

Finally, in the field of HCI, compared, for instance, to the field of education, activity theory is not yet operationalized enough. There are not enough methods and techniques that can be directly utilized to solve specific problems, so it would be unrealistic to expect immediate results from accepting activity theory as an approach guiding theoretical research or practical efforts.

The limitations of activity theory set out are not inevitable. It is a developing approach, and probably one of its strengths is its potential for integration with other conceptual systems.

ACKNOWLEDGMENTS

I am grateful to Rachel Bellamy, Susanne Bødker, and Bonnie Nardi for comments on prior drafts of this chapter.

NOTES

1. Usually cognitive scientists avoid explicit identification of the mind with the computer. It is stated that the basic concept underlying the cognitive approach is the concept of computability: ``Theories of the mind should be modelled in a computer program" (Johnson-Laird 1988). It does not necessarily mean that the human mind is a computational device. However, concrete cognitive studies of the mind are based on the explicit or implicit assumption that the nature of the human mind is information processing, and it can be described in terms of architecture, procedures, flow of information, distributed processing, and so forth (Anderson 1983; Gardner 1987; Johnson-Laird 1988).

2. This chapter does not discuss the limitations of GOMS that can potentially be overcome within the cognitive approach, such as its inability to deal with parallel activities or its overreliance on the average times taken by elementary operations.

REFERENCES

Allwood, C. M. (1989). Computer usage by novices. In I. A. Kent and J. G. Williams, eds., *Encyclopedia of Microcomputers*. New York: Marcal Dekker.

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Bannon, L. (1991). From human factors to human actors: The role of psychology and human computer interaction studies in system design. In J. Greenbaum and M. Kyng, eds., *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, NJ: Lawrence Erlbaum.

Bateson, G. (1972). *Steps to an Ecology of Mind*. New York: Ballantine Books.

Bødker, S. (1989). A human activity approach to user interfaces. *Human Computer Interaction* 4(3):171– 195.

Bødker, S. (1991). *Through the Interface: A Human Activity Approach to User Interface Design*. Hillsdale, NJ: Lawrence Erlbaum.

Bowers, J., and Rodden, T. (1993). Exploding the interface: Experiences of a CSCW network. In *Proceedings INTERCHI'93 Conference on Human Factors in Computing Systems* (pp. 255–262). Amsterdam, April 24–29.

Brusilovsky, P., Burmistrov, I., and Kaptelinin, V. (1993). Structuring the field of HCI: An empirical study of experts' representations. In *Proceedings East-West Human-Computer Interaction Conference* (vol. 3, pp. 18–28). Moscow, August 3–7.

Card, S., Moran, T., and Newell, A. (1983). *The Psychology of Human Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum.

Carroll, J. M., Kellogg, W. A., and Rosson M. B. (1991). The task-artifact cycle. In J. Carroll, ed., *Designing Interaction: Psychology at the Human-Computer Interface*. Cambridge: Cambridge University Press.

Cole, M. (1984). The zone of proximal development: Where culture and cognition create each other. In J. Wertsch, ed., *Culture, Communication, and Cognition: Vygotskian Perspectives*. Cambridge: Cambridge University Press.

Cypher A. (1986). The structure of users' activities. In D. Norman and S. Draper, eds., *User Centered System Design: New Perspectives on Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum.

Engeström, Y. (1987). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Helsinki: Orienta-Konsultit Oy.

Gardner, H. (1987). *The Mind's New Science: A History of the Cognitive Revolution*. New York: Basic Books.

Grudin, J. (1989). The case against user interface consistency. *Communications of the ACM* 32(10):1164–1173.

Grudin, J. (1990). The computer reaches out: The historical continuity of interface design. In *Proceedings CHI'90 Conference on Human Factors in Computing Systems* (pp. 261–268). Seattle, Washington, April 1–4.

Grudin, J. (1993). Interface: An evolving concept. *Communications of the ACM* 36(4):110–119.

Johnson-Laird, P. (1988). *The Computer and the Mind: An Introduction to Cognitive Science*. Cambridge, MA: Harvard University Press.

Kaptelinin, V. (1992a). Integration of computer tools into the structure of human activity: Implications for cognitive ergonomics. In *Proceedings Sixth European Conference on Cognitive Ergonomics* (pp. 285– 294). Balatonfured, Hungary, September 6–11.

Kaptelinin, V. (1992b). Human computer interaction in context: The activity theory perspective. In *Proceedings East-West Human Computer Interaction Conference* (pp. 13–28). St. Petersburg, Russia, August 4–8.

Kaptelinin, V. (1993). Item recognition in menu selection: The effect of practice. In *Adjunct Proceedings, INTERCHI'93 Conference on Human Factors in Computing Systems* (pp. 183–184). Amsterdam, April 24–29.

Kuutti, K., and Bannon, L. (1993). Searching for unity among diversity: Exploring the interface concept. In *Proceedings INTERCHI'93 Conference on Human Factors in Computing Systems* (pp. 263–268). Amsterdam, April 24–29.

Latour, B. (1993). On technical mediation. Three talks prepared for ``The Messenger Lectures on the Evolution of Civilization." Cornell University, April.

Leont'ev, A. N. (1978). *Activity. Consciousness. Personality*. Englewood Cliffs, NJ: Prentice-Hall.

Leont'ev, A. N. (1981). *Problems of the Development of Mind*. Moscow: Progress.

Mayes, J. T., Draper, S. W., McGregor, M. A., and Oatley, K. (1988). Information flow in a user interface: The effect of experience and context on the recall of MacWrite screens. In D. M. Jones and R. Winder, eds., *People and Computers IV*. Cambridge: Cambridge University Press.

Miller, G. A., Galanter, E., and Pribram, K. (1960). *Plans and the Structure of Behavior*. New York: Holt, Rinehart, and Winston.

Monk, A., Nardi, B., Gilbert, N., Mantei, M., and McCarthy, J. (1993). Mixing oil and water? Ethnography versus experimental psychology in the study of computer-mediated communication. In *Proceedings INTERCHI'93 Conference on Human Factors in Computing Systems* (pp. 3–6). Amsterdam, April 24–29.

Moran, T. (1981). The Command Language Grammar: A representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies* 15.

Munipov V. M. (1983). Leont'ev's contribution to engineering psychology and ergonomics [Vklad A. N. Leontjeva v razvitie inzhenernoj psikhologii i ergonomiki]. In A. V. Zaporozhets, V. P. Zinchenko, and O. V. Ovchinnikova, eds., *Leont'ev and Contemporary Psychology* [A. N. Leontjev i sovremennaja psikhologija]. Moscow: Izdatelstvo MGU. (in Russian)

Norman, D. (1988). *The Psychology of Everyday Things*. New York: Basic Books.

Norman, D. (1991). Cognitive artifacts. In J. Carroll, ed., *Designing Interaction: Psychology at the Human-Computer Interface*. Cambridge: Cambridge University Press.

Norman D. (1993). *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Reading, MA: Addison-Wesley.

Petrovsky, A. V., and Petrovsky, V. A. (1983). Active personality and Leont'ev's ideas [Lichnost i ee aktivnost v svete idej A. N. Leontjeva]. In A. V. Zaporozhets, V. P. Zinchenko, and O. V. Ovchinnikova, eds., *Leont'ev and Contemporary Psychology* [A. N. Leontjev i sovremennaja psikhologija]. Moscow: Izdatelstvo MGU. (in Russian)

Ponomarev Ja. A. (1975). *The Psychology of Creativity* [Psikhologija Tvorchestva]. Moscow: Nauka. (in Russian)

Raeithel, A. (1992). Activity theory as a foundation for design. In C. Floyd et al., eds., *Software Development and Reality Construction*. Berlin: Springer Verlag.

Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. (1993). The cost structure of sensemaking. In *Proceedings INTERCHI'93 Conference on Human Factors in Computing Systems* (pp. 269–276). Amsterdam, April 24–29.

Shneiderman, B. (1987). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Reading, MA: Addison-Wesley.

Suchman L. (1993). Centers of coordination: A case and some themes. Presented at the NATO ARW on Discourse, Tools, and Reasoning, Lucca, Italy, November 2–7.

van der Veer, G. (1994). Groupware task analysis: Modeling complex reality. In *Proceedings Seventh European Conference on Cognitive Ergonomics* (pp. 341–352). Bonn, Germany, September 5–8.

Vygotsky, L. S. (1978). *Mind and Society*. Cambridge MA: Harvard University Press.

Zinchenko, V. P. (1992). Activity theory: Retrospect and prospect. In *Proceedings East-West Human Computer Interaction Conference* (pp. 1–5). St. Petersburg, Russia, August 4–8.

Figure 3.1
Expansion of the subject matter of HCI: three dimensions.

Figure 3.2
Computer tool as an extension of the internal plane of actions (IPA).

Figure 3.3
Computer-mediated group activity. ccs = central cognitive structures; IPA = internal plane of actions; ind. art. = individual artifact; grp. art. = group artifact.