2000

INTUITIVE USER INTERFACES

Klaus B. Bærentsen

Bang & Olufsen, Denmark, klaus@psy.au.dk

Follow this and additional works at: http://aisel.aisnet.org/sjis

INTUITIVE USER INTERFACES

Klaus B. Bærentsen

Bang & Olufsen A/S. Dept. 3700-1, DK-7600 Struer, Denmark.
Now at MR-Research Center and Institute of Psychology, University of Aarhus, Denmark.
E-mail: klaus@psy.au.dk

Abstract

The paper outlines an approach to the development of intuitively understandable on-screen user interfaces. Users have been found to explain the operation of equipment with screen based user interfaces in terms of handling "objects" and interacting with "agents" in a virtual "space". These metaphorical descriptions may reflect general and fundamental principles of cognition that are rooted in the evolution of the human species. It is postulated that presentation of information on the interface as scenes, objects and actors can call upon instinctive capacities for direct perceptual information pickup, intuitive cognitive functions and natural behavioral tendencies. In order to initiate learning of complex functions that cannot be perceived directly may necessitate the use of symbolic information. This must be based on an analysis of the most appropriate way to map the new functions to the users' prior conceptual understanding of technological objects and functions.

Keywords:

Activity theory, ecological psychology, intuition, user interface, objects, scenes, actors

INTRODUCTION

Mediating experience

Mediation of experience by technological means has a long history. Sculptures, paintings, and other devices have been utilized for recording, writing, calculation and transmission of information for thousands of years. Binoculars, telescopes, and microscopes were invented hundreds of years ago. The telegraph, telephone, gramophone, radio and television are more recent inventions. Today, computers with complex programmed functions reside in lots of products used to mediate such activities.
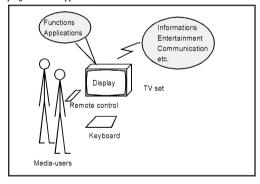
In mechanical devices the functions and their control structure is objectified in the material structure of the mechanics and thus have a perceptible surface. In contrast to this, functions in programmed technology have no inherent perceptual form, and the user interface is made as an explicit design, based on the imagination of the designer, and constrained by tradition and the possibilities and limits of the programming language and screen technology.

Interface technology has evolved rapidly since the first computers entered the consumer market. Command based interaction evolved into 2D choice menus and graphical interfaces utilizing the metaphor of the "desktop" (e.g. Windows) (see e.g. Shneiderman 1983, Myers 1998, Perkins, Smith Keller & Ludolph 1997). Recently some researchers and designers suggest the use of 3D graphical interfaces for advanced applications (Robertson 1999). In this paper it will be argued, that a logical step forward would be to think of 4D interfaces, utilizing the metaphors of objects, scenes and agents.

State of the art user interfaces

Although "Windows" has not conquered the area of consumer goods, it is justified to say that the evolution of interfaces for e.g. televisions have passed through similar stages since the invention of the "remote control" and the removal of hardware buttons and inter faces from the "corpus" of the TV. Today most TV's come with on-screen user interface (OSD) in the form of menus designed to enable the user to control the functions and applications in the TV set, to select different sources of entertainment and information via the remote control, to key in necessary data to the systems functions, and thereby to manipulate features and options determining the appearance of pictures on the screen (see Figure 1).

Figure 1: Media users access the world of information, entertainment and communication via functions and application programs in an apparatus.

Different brands have different philosophies concerning the construction and appearance of such on-screen user interfaces. At Bang & Olufsen A/S (B&O) [1] the aim is to provide the best of two worlds, aesthetic-emotional experiences and technical functions. B&O products should enable the user to access and enjoy the cultural experiences available via various media in a direct and pleasurable way. This calls for design of "transparent" and aesthetically pleasing user interfaces

by multidisciplinary teams, as recently described by Bærentsen & Slavensky (1999). A "transparent" user interface should not be invisible however, but noticed and even appreciated for this transparency. From a usability point of view the challenge is consequently to make products so easy to operate, that the operation will become an enjoyable experience in itself.

Most user interfaces today present system functions by use of verbal or iconic symbols on static 2D menu pages organized in a hierarchical system. From the menus the user can select and activate functions and additional menus by pressing a button on a keyboard, a remote control or the like. Activation of a function may be accompanied by verbal status information giving "feedback" about the effects. When the user selects a menu (or a submenu) it appears instantly, replacing the previous menu in an abrupt manner. In order to orient the user about his whereabouts in the system each menu usually has a headline indicating its contents.

Many variations of details exist, but the similarities are dominating. Even though in graphical interfaces for computers some functions are handled in a non-symbolic way (e.g. "drag and drop") they are mostly supported by a symbolic "dialogue" requesting confirmation of the intended operation.

It is thus mainly the "highest" and most specific intellectual capabilities of humans that are taken into account in current user interfaces, i.e., the abilities for reading and understanding linguistically symbolized information, for logical reasoning, and conscious decision making. Only in computer games designers have gone beyond the limits of symbolic information processing - (linguistic and/or iconic) - and into the realm of basic sensory-motor and perceptual processes.

All users of advanced audiovisual equipment have some previously acquired capabilities for using technology that may be more or less relevant to the operation of new or unknown equipment. Some of these background capabilities are specific for the individual because of individual experience, professional training or as a member of a given culture. But all users are members of the human species, and as such they share a number of characteristic competencies. Some of these qualities are general features of the human species, and some of them are not even specific to humans, but have a distant evolutionary background and are shared with a broad spectrum of living creatures. In order to tap these competencies it is necessary to take a look at their functional origin and nature, and the kinds of information that they are dependent on.

The human brain did not evolve primarily in order to handle symbolic information processing. Judged by the archaeological record the use of externally objectified symbols probably emerged some 50-30.000 years a go, and writing originated only 5-6.000 years ago (Donald 1991, Klix 1993; Schmandt-Besserat 1978). The human brain reached its contemporary form approximately 200.000-100.000 years ago. Some of its essential qualities are as old as the transition of life from the water to the solid ground several hundred million years ago, and some capabilities are even older than that. The primary function of the brain during this lengthy period was the non-symbolic control of locomotion, object manipulation and interaction with other living creatures.

INTUITIVE USER INTERFACES

In the following an approach to the development and construction of intuitively understandable on-screen user interfaces will be sketched. First a definition of intuition and intuitive interfaces will be introduced. Then some empirical findings of metaphors employed by users to explain their

understanding of computer systems and functions of a TV set will be presented. These metaphors point to fundamental principles of human activity, and its control by hierarchically organized systems of perception-action and cognitive functions rooted in the evolution of the human species. Some of these principles are then summarized on the basis of activity theory, cultural historical psychology, ecological psychology, and the theory of situated actions. Later some suggestions are given on how to implement these principles in interface technology.

Intuition

According to a standard dictionary (Oxford 1989, p. 660) intuition is the "(power of) understanding things ... immediately, without the need for conscious reasoning or study". Etymologically "intuition" is derived from Latin: *intueri*, look into, observe (Brüel & Nielsen 1987, p. 262). In a psychological encyclopedia intuition is described as being "... essentially arriving at decisions or conclusions without explicit or conscious processes of reasoned thinking. It is sometimes thought that intuitions are reliable; and indeed we do act most of the time without knowing why or what our reasons may be. It is certainly rare to set out an argument in formal terms, and go through the steps as prescribed by logicians. In this sense, almost all judgments and behavior are 'intuitive'" (Gregory 1989 p. 389).

Definition of an intuitive interface

Hence an *intuitive interface* may be defined as an interface, which *is immediately understandable to all users, without the need neither for special knowledge by the user nor for the initiation of special educational measures. Anybody can walk up to the system; see what kind of services it affords, and what should be done in order to operate it. While operating the device, navigation and manipulation of the system interface should proceed without the need for conscious awareness of the sensory- motor operational aspects of the interface*.

   Since this definition presupposes an initial knowledge by the user of the kinds of functions embodied in the technology, this degree of intuitivity is probably not achievable. Indeed it will be *unattainable in principle*, inasmuch as the development of information technology is very fast. It is therefore necessary to supplement the definition of an intuitive interface with *the availability of functions supporting learning of unknown functions and their operation, but in a way that is not perceived as "teaching" or "education". Learning must be a spontaneous product of the activity of use*.

Levels of an intuitive interface

It needs to be specified that not all aspects of the use activity are considered in this paper: for instance the motivational-intentional aspects of interaction related to the practical relevance and conceptual understanding of the functions of the system in question. In order for a system to be operated "intuitively", the user of course needs to have a motive for using the system, and at least a minimal understanding of its affordances (i.e. the functions it affords). No matter how sophisticated the interface is, the system cannot be "operated" in an intuitive way, if the functions are completely useless and unknown to the user. Thus it is presupposed that the user has appropriated a suitable motivation and a general understanding of the meaningfulness of functions in the

system.

The theme of this paper is sensory-motor aspects of interaction, that today are inappropriately delegated to control by conscious intellectual, linguistic and symbolic functions, that would be better served by direct and automatic perception- action capabilities outside the field of conscious attention. Some examples from empirical investigations will be given in the next section.

EMPIRICAL DATA ON USERS' UNDERSTANDING OF SYSTEMS

Metaphorical descriptions of use activities

In an empirical investigation of librarians using a computerized card index system (Andersen & Madsen 1989) and an investigation of people making programmed recordings on VCR's with an on-screen menu system (OSD) (Bærentsen et al. 1995) it was found that the users described their activities using expressions like "now I'm going into ...", "now I'm going out of ..., - then I go up to ..., - then I go down to ..." etc. and "it tells me to", "then it remembers" or "it didn't do what I told it to".

The utilization of these metaphorical forms of expression suggests that the user's activity to-wards equipment with screen-based user interfaces is conceivable as activity in a virtual "spatial environment", handling and manipulation of virtual "objects" and interaction with virtual "actors". The mental models (Bedny & Meister 1997 pp. 14 ff., 96 ff., 131 ff.; Norman 1988, pp. 12 ff.) which contains the user's understanding of the system, and serves to regulate the actions, are drawing on analogical similarities (Carroll & Thomas 1982; Carroll & Mack 1984; Holyoak & Thagard 1997) to normal actions with real objects in the physical environment. The user's intu-itions about what actions would likely result in a particular outcome, the expectations as to what will happen when some action is taken, or what might be the reasons leading to a particular situ-ation or result, is to a large degree informed by experience from analogical daily activities. In the following section the results from research done at Bang & Olufsen will be presented in order to explicate some of these principles.

"Objects" on the interface of TV's

In a usability test carried out at B&O we found indications for the validity of the metaphor "object" as a description of some aspects of the mental models regulating of the users actions. The test was carried out during the late stages of the development of a recent B&O TV, the "AV5" integrated AV set (see Figure 3). The user interface on the AV5 was developed by a dedicated designer, and introduced a new generation of interfaces replacing earlier types of software-developer generated interfaces present on the Avant TV (see Figure 2) and earlier B&O TV's. The design of the interface for the AV5 was inspired by an interface developed by Ideo for a Nokia TV. It was therefore con-sidered relevant to set up a comparative test involving the Avant, the AV5 and the Nokia TV.

The test verified that the AV5 menu interface supported the user interaction in a uniquely pleas-ant way, and indicated some of the principal shortcomings of the interface of the Avant set. The test also verified and clarified the understanding of the principles and factors determining the user's ease of operation and understanding of the process. The test verified that the AV5 menu

Klaus B. Bærentsen

interface supported the user interaction in a uniquely pleasant way, and indicated some of the principal shortcomings of the interface of the Avant set. The test also verified and clarified the understanding of the principles and factors determining the user's ease of operation and understanding of the process.

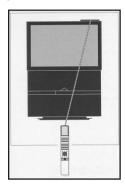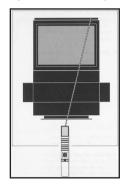**Figure 2: The B&O Avant TV.**

Figure 3: The B&O AV5 integrated AV set.

One example of this relate to the way TV channels are stored when they are tuned in and the resulting order of appearance of the channels on the channel list (see Figure 4 and Figure 5).

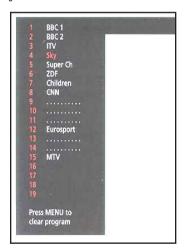**Figure 4: The Channel list on the Avant TV.**

Figure 5: The Channel list on the AV5 integrated AV set.

In Denmark many or most TV sets are connected to cable networks on which the particular cable network operator determines the transmission frequency of the individual channels. On one of the major cable net works the Danish national television channel "DR1", which is considered to be "the first channel" by most Danes, at the time of the test was transmitted on 062 MHz, whereas the competing "TV2" - which logically enough is considered the "second channel" by most Danes - was transmitted at 046 MHz. During the tuning process "TV2" consequently appeared first, where-

34

Klaus B. Bærentsen

as "DR1" appeared as the second channel.

After conclusion of the tuning process, most users would therefore wish to edit the channel list in order to have DR1 become channel one, and TV 2 channel two. DR1 would then be presented when pressing the "1" button on the BEO4 terminal, and TV2 would appear when button "2" is depressed (see Figure 6). This of course seems more logical than the opposite.

On the Avant the tuning process has to be accomplished semi-manually by finding, numbering, naming and storing each channel one at a time. In order to do this, the user must first call up the main menu (Figure 7), then activate the setup menu (Figure 8), the tuning menu (Figure 9), and then the TV tuning menu (Figure 10) where the tuning can be accomplished.

Figure 6: The Beo4 terminal.

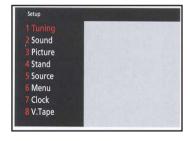*Figure 7: Avant main menu.*

Figure 8: Avant setup menu.

Figure 9: Avant tuning menu.

Figure 10: Avant TV tuning menu.

Klaus B. Bærentsen

35

After completion of the tuning, the order of channels on the list may be edited (see Figure 4). In order to do this the user has to select the channel to be moved, and again call up the menus (Figure 7 to Figure 10), and copy the channel to a vacant position in the channel list (i.e. set up a dummy variable), by assigning a new (unused) channel number to it. Then - using the same procedure - the other channel that is to be moved can be copied to the position of the first channel, overwriting this. In the last step, by repeating the procedure, the channel that is stored in the dummy variable can be copied to the initial position of the second channel, overwriting that one. The user can then utilize the dummy variable for more swapping, or just delete the channel that was temporarily stored there. The complete procedure for the interchange of positions of "Channel_1" and "Channel_2" on the B&O Avant TV is illustrated in table 1.

In the table, # denotes a number, *Channel_1* and *Channel_2* are the channels that are interchanged, after which they become *Channel_1** and *Channel_2**. Text in square brackets [] denotes activation of a button on the Beo4 terminal.

On the AV5 set, the tuning can be accomplished automatically by activating the "Autotune" function on the tuning menu (see Figure 11). The user only has to wait while the set finds, names and stores the channels. When the auto tuning process is completed, the channel list is presented, and the user may then name those channels that could not be named automatically, and edit the order of the channels on the list (see Figure 5 and Figure 12).

Figure 11: Tuning menu on the AV5 integrated AV set.

In order to exchange the position of two channels on the list, the user places the cursor on the name of the channel to be moved, activates the "move" function, which transfers the channel name to the free space on the right side of the channel list, and can then move the channel up or down the list (see Figure 12).

Figure 12: Changing the position of a TV channel on the AV5

Table 1: Procedure for editing the channel list on the B&O Avant TV

SITUATION	USER ACTION ON BEO4	OUTCOME
SUBTASK: FIND *CHANNEL_1*		
TV is on	[MENU]	Main Menu
Main Menu	[3 = TV list]	TV channel list
TV channel list	[# of *Channel_1*]	Cursor on # *Channel_1*
Cursor on # *Channel_1*	[GO]	TV on *Channel_1*
SUBTASK: COPY *CHANNEL_1* TO *DUMMYCHANNEL*		
TV on *Channel_1*	[MENU]	Main Menu
Main Menu	[4 = Setup]	Setup Menu
Setup Menu	[1 = Tuning]	Tuning menu
Tuning menu	[1 = TV]	TV tuning menu
TV tuning menu	[▶▶]	Cursor on # of *Channel_1*
Cursor on # of *Channel_1*	[# of *DummyChannel*]	Cursor on # of *DummyChannel*
Cursor on # of *DummyChannel*	[GO]	Tuning menu (*Channel_1* copied to *DummyChannel*)
Tuning menu	[EXIT]	TV on *DummyChannel*
SUBTASK: FIND *CHANNEL_2*		
TV on *DummyChannel*	[MENU]	Main Menu
Main Menu	[3 = TV list]	TV channel list
TV channel list	[# of *Channel_2*]	Cursor on # of *Channel_2*
Cursor on # of *Channel_2*	[GO]	TV on *Channel_2*
SUBTASK: MOVE *CHANNEL_2*		
TV on *Channel_2*	[MENU]	Main Menu
Main Menu	[4 = Setup]	Setup Menu
Setup Menu	[1 = Tuning]	Tuning menu
Tuning menu	[1 = TV]	TV tuning menu
TV tuning menu	[▶▶]	Cursor on # of *Channel_2*
Cursor on # of *Channel_2*	[# of *Channel_1*]	Cursor on # of *Channel_1*
Cursor on # of *Channel_1*	[GO]	Tuning menu (*Channel_2* copied to former position of *Channel_1* ? *Channel_2**)
Tuning menu	[EXIT]	TV on *Channel_2**
SUBTASK: MOVE *DUMMYCHANNEL*		
TV on *Channel_2**	[# of *DummyChannel*]	TV on *DummyChannel*
TV on *DummyChannel*	[MENU]	Main Menu
Main Menu	[4 = Setup]	Setup Menu
Setup Menu	[1 = Tuning]	Tuning menu
Tuning menu	[1 = TV]	TV tuning menu
TV tuning menu	[▶▶]	Cursor on # of *DumyChannel*
Cursor on # of *DummyChannel*	[# of *Channel_2*]	Cursor on # of *Channel_2*
Cursor on # of *Channel_2*	[GO]	Tuning menu (*DummyChannel* copied to former position of *Channel_2* ? *Channel_1**)
Tuning menu	[EXIT]	TV on *Channel_1**

Klaus B. Bærentsen

37

When it is positioned to the right of the desired position, the user presses the button for the "swap", and the two concomitant channels exchange their positions. In order to complete the process, the user then moves the "active" channel to the desired position, and again press the button assigned to the "swap" function. The complete procedure for the interchange of positions for "Channel_1" and "Channel_2" on the B&O integrated AV system AV5 is illustrated in table 2:

Table 2: Procedure for editing the channel list on the B&O AV5 integrated AV set

SITUATION	USER ACTION	OUTCOME
SUBTASK: FIND CHANNEL_1		
TV	[MENU]	Setup Menu
Setup Menu	[GO]	TV setup menu
TV setup menu	[▼]	TV setup menu, cursor on PROGRAM LIST
TV setup menu, cursor on PROGRAM LIST	[GO]	Program List
Program List	[# of Channel_1]	Cursor on Channel_1
SUBTASK: MOVE CHANNEL_1		
Cursor on Channel_1	[▶▶]	Cursor on Channel_1 in right column
Cursor on Channel_1 in right column	n* [▲] or [▼]	Cursor on Channel_1 in right column next to Channel_2 in left column
Cursor on Channel_1 in right column next to Channel_2 in left column	[◀◀]	Cursor on Channel_2 in right column next to Channel_1* in left column
SUBTASK: MOVE CHANNEL_2		
Cursor on Channel_2 in right column	n* [▲] or [▼]	Cursor on Channel_2 in right column next to empty former position of Channel_1 in left column
Cursor on Channel_2 in right column next to empty former position of Channel_1 in left column	[◀◀]	Cursor on Channel_2* in left column
Cursor on Channel_2* in left column	[EXIT]	TV on Channel_2*

It is obvious that the two methods differ very much from each other. The test demonstrated that many people were unable to find a solution to the problem of changing the position of channels on the Avant, and instead proposed to tune the set again, and this time store the channels in the proper positions. Those who found the solution however had a hard time to complete the procedure. Even those (mainly engineers and programmers) who immediately found the solution, very soon ran into problems when they were requested to continue sorting the channels on the list. After a few repetitions it was very difficult to remember which channel they were about to move, which one was stored in the dummy variable, and which one had already been moved. Likewise it was cumbersome to remember the procedure - go to the channel that should be moved, call up the tuning menu, change the number of the channel, store the channel, exit tuning etc.

In contrast to this, the only major problem encountered on the AV5 set was to notice the text on the screen menu indicating which button to press on the terminal in order to initiate the process. The rest of the procedure seemed to be readily "graspable", and was accomplished without the need for much attention to the process, and without any major problems.

Errors and misunderstandings, or problems with the interface?

When users were grappling with the tasks of moving the TV channels around in the channel list on the Avant, some typical forms of errors were encountered.

For example some users attempted to swap channels directly, by changing the number of the channel they wished to move to the wanted number, without copying the channel to a dummy variable first (e.g. instead of changing "48, *1*, TV2" to "48, *dummy*#, TV2", they directly changed to "48, *2*, TV2"; see Figure10). When they were asked what they thought had happened to the channel that was originally stored on position "2", they first assumed it was saved somewhere "behind" in a kind of buffer, but then realized they had made a mistake and deleted the channel.

Other users tried to change the name of the channel into the name of the desired channel (e.g. "48, 1, *TV2*" was changed to "48, 1, *DR1*"). The expectation was then that "DR1" would appear when they chose channel 1. When the picture of "TV2" appeared instead, but with the name "DR1", they realized the assumption was in error, and that they had only changed the name of the channel.

On the AV5 some users encountered one minor problem in the procedure when the channel to be moved was placed alongside the channel occupying the position it should be moved to. When the "Swap" function is activated, the swap is accomplished instantly, i.e. the names disappear, and reappear fractions of a second later in the opposite positions. In order to establish that the channels have changed position, the user must read the names of the channels, identify them and compare their positions to the remembered prior situation and note the difference. This demands the user's conscious attention to the channels and application of logical symbolic reasoning. Later it was shown that if the swap is instead done by continuous "sliding" of the names behind each other, the exchange can be perceived directly, without the user having to attend consciously to the process.

Although the Avant was still difficult to operate for those who were familiar with the nature of programmed devices, the outcomes of actions were not impossible to understand. The "errors" and problems encountered by users without this special knowledge, i.e. without a training in programming are easily explicable if it is assumed that they conceived the representations of the TV channels (the picture on the screen, the transmission frequency, the number of the channel on the list, and the name of the channel) as features belonging to physical objects with a kind of substantial identity, whereas in fact they are only parameter values for data structures in a program

The empirical data support the hypothesis that the user's understandings and intuitions about the behavior of the system are built by drawing analogies from the everyday experience of handling physical objects in the physical world. If this seems surprising at first sight, after a short reflection it will appear as a rather obvious insight. When confronted with something new, people will employ whatever experience they may have in their efforts to understand new phenomena. And when they encounter some things that appear to be objects, they will intuitively expect them to behave like objects.

To the "uneducated" user the "mental model" of a TV channel can be likened to that of an object with a number of features, like a book on a bookshelf. A book has a meaningful content, a title, form, size and color, and a physical substance, and it occupies a particular position on the shelf. If the book is moved to another position on the shelf, it is not only the content, but also the title

of the book, its form, color etc. that move. Since the substance moves as well, any other book that may already be on the position to which we move it, will be displaced, but not annihilated by the newcomer. In the "real" world, the world of physics, objects have a basic *substantial quality* or *numerical identity* (see Mammen 1989, 1993, 1994; Mogensen 1997; Xu & Carey 1996). This fundamental substantiality dictates certain kinds of behavior and precludes others.

But when users enter the world of programmed technology, like the poor Alice they enter a kind of "wonderland" in which the "laws of nature" no longer apply. And their expectations are not fulfilled. As in a cartoon anything may happen. A stored TV channel is a data structure with a number of abstract variables of certain types, to which some arbitrary (although valid) values can be assigned. When the user assigns values to the variables, the resulting set of data (e.g. the frequency on the transmission band, a digit, a string of characters, and the contents of the signal) can be presented simultaneously on the screen, but they are an arbitrary conglomeration of distinct abstract entities without internal coherence, and the rules governing their combination are purely conventional. In the world of programmed technology "objects" are conventional concatenations of *abstract variables* with assigned values, and have no substantial existence.

The shortcomings of the Avant interface consequently stems primarily from its almost complete reliance on the users ability to perform conscious reasoning on a symbolic information processing level. The user has to read, understand and reason verbally in order to understand the abstract information on the interface. It even requires logical thinking and rational problem solving in order to figure out how to accomplish some tasks.

On the AV5 and the Nokia sets some of these tasks are accomplished by direct manipulation of "objects" on the screen. This transfers the users thinking from the level of conscious symbolic reasoning to the levels of automatic perception-action and sensory-motor functions, which makes it much easier to carry out the necessary operations. On the Nokia set the direct manipulation of objects was achieved by "imitation" of "real objects" on the screen. In this way some crucial principles were "imported" into the interaction. On the AV5 the features supporting the principles of direct manipulation of objects were maintained, but the nonessential imitational features were skipped. This turned out to be an advantage because some distracting visual features in this way disappeared from the screen.

Systems as "Spatial environments"

The discontinuous way current hierarchical menu systems present navigation in the functional space that is represented on the menus pose comparable problems to the user. The instantaneous shifts from one menu page to the next require the user to know the architecture of the system, or to think consciously about his movements in the system, or to read and understand the "status information" on the screen in order to keep oriented about his position in the menu hierarchy.

This problem was very obvious in a power plant control room having a computerized control system with alphanumeric displays (cf. Bærentsen 1992; Bærentsen, Kvorning & Skov 1989). When the activity of the operators was analyzed, it turned out that they had to consciously control the presentation of information on the interface in order to keep oriented about what part of the system was represented and controlled from the particular display. The control of the interface had become a particular task on top of the main tasks of controlling the processes, the machine sys-

tems, and the automatic control system.

As with objects, if we compare navigation in the functional spaces of programmed devices with the way we navigate in the real world, the difference is one of substantiality and continuity. Instantaneous jumps from one location to another are simply impossible in the real world, and when we move from one place to another, the scenery changes continuously and in a way that automatically and unambiguously informs us about our whereabouts during the transition. We rarely need a sign giving "feedback" information about our position.

The use of metaphors in the description of activities with the technical devices is not a singular phenomenon or one specific to technology. Metaphors are ubiquitous in human description of activities and phenomena of life (Johnson 1987, Lakoff 1987). As these authors indicate it should not be interpreted as a purely linguistic phenomenon, but rather be taken as a serious indication of the kind of imagistic mental processes involved in reasoning and talking about processes and phenomena in the world (Barsalou 1999, Chafe 1990, Damasio 1989; Deecke 1996, Farah 1984, Jeannerod 1994, Larsen 1989). What is hypothesized in this paper is that these findings can be generalized to the domain of human activities with technology (cf. Bærentsen 1996).

On this basis some general principles can be derived for the presentation of information and interaction with functions, which are supported by automatically operating natural capacities for perceptual pickup of information, intuitive cognitive functions and behavioral tendencies. These principles start out from a consideration of the phenomena implied at in the metaphors of scenes, objects and actors.

THE PSYCHOLOGICAL BASIS FOR INTUITION

Psychological theories

As mentioned above, the assertions made in this paper are based on sources from cultural historical psychology (e.g. Vygotsky 1966), activity theory (e.g. Bedny & Meister 1997; Bernshtein 1996; Leont'ev 1977, 1978, 1979, 1981; Leont'yev et al. 1966a, 1966b; Lomov 1963; Rubinshtein 1934; Velichkovsky 1994; Zinchenko & Munipov 1989), ecological psychology (e.g. Chan & Shaw 1996; Effken, Kim & Shaw 1997; Gibson 1966, 1986; Reed 1996; Turvey 1990; Turvey & Carello 1986; Vicente & Rasmussen 1992), and the theory of situated actions (Suchman 1987). These theories have different historical origins, address different questions, and differ in many details. Anyway they also share fundamental viewpoints on human activity, and they fit very well together when it comes to their practical consequences in relation to the theme of the present effort.

According to the theory of situated actions, user's intentional activities (i.e. their operational realization) with technology are not determined solely by the intention or a "cognitive plan", but depends to a large degree on the contextual circum stances of the actions, hence the notion "situated actions". Unfortunately however the contextual determination is emphasized to such a degree, that the real influence of motives, intentions, goals and plans as structuring factors for actions are nearly lost (see Bardram 1997).

The core of Suchmans argument fits nicely with the insight of activity theory, that it is necessary to differentiate three aspects of activity: the motivational determination, the determination by conscious goals and intentions, and the determination by the situational means and conditions

Klaus B. Bærentsen

(Leont'ev 1979). The operational realization of motivated and goal-directed activity as a process in the four dimensional space of the physical world requires adaptation of the actions to the contextual conditions, but this does not do away with the determining influence of neither motives nor intentions (see Figure 13).
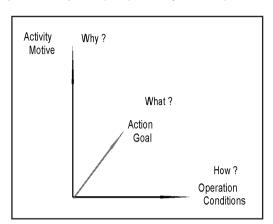
Figure 13: According to activity theory different aspects of activity are controlled by different factors.

One advantage of activity theory over the theory of "situated actions" and the ecological theory is a differentiated concept of activity, with a clear appreciation of the distinctions and relations between the emotional-motivational, the intentional and the concrete sensory-motor operations realizing the activity of the individual. This is a clear advantage in regard to understanding the different sources of information for different aspects of cognitive and sensory-motor control of action.

Evolution of the brain

In activity theory the different "layers" or "dimensions" of determination are discerned by an analysis of the phylogenetically established "morphology" of the sensory-motor aspect of activity and its control. This is summarized in the following paragraphs on the basis of a free interpretation of data and theory in (Arbib 1989; Bernstein 1996; Fabri 1983; Gibson 1966, 1986; Klix 1993; Leont'ev 1978; Leontyev 1977, 1978, 1979, 1981; Neisser 1994; Velichkovsky 1994). The summary is only tentative, and makes no claim of being neither exhaustive nor accurate in detail. Neither are any exact references to the literature given by which the specific contributions of the separate sources can be traced.

Life activity

The most fundamental characteristic of living organisms is their *activity*; no matter how primitive or complex, they are inherently active. This is especially true for animals that move around in their surroundings actively searching for substances that may serve to satisfy their needs. This motivated activity is also the real basis of our knowledge of the world. In relation to the use of

technology this implies that the use of devices serves to satisfy some kind of need on the part of the user, and that the fundamental trick in relation to the user's learning of the device must be to get the user to *do something* with the equipment, and that the resulting effects should provide the user with information about the functions of the device.

Locomotion and object manipulation

All vertebrates share some general characteristics in the way they move about in the world. Humans in particular have many features in common with the land- living vertebrates with whom we share the ecological conditions of locomotion, and consequently the general architecture of our skeleton and nervous system (including our brain and sensory organs). The ability for locomotion through the environment is served by a capacity for dynamic sensory localization of objects in relation to the organism and in relation to other objects. The aim of much of this locomotion is the achievement of need related objects, and this entails the ability to perceptually identify encountered objects as instances of natural classes of objects with certain affordances (i.e. qualities affording certain activities).

Although rodents and some other animals use their forelimbs to grasp and handle objects, extensive use of the hands for "manipulation" of objects is first seen in hominids, where this kind of activity attains a highly significant role. Together with the ability to conceptualize functional affordances of objects making them suitable as instruments in relation to the achievement of complex goals this constitutes an ability for "manual thinking" in concrete situations. This kind of thinking is constrained by the need for concrete objects, either physical or imagined, and by the absence of generative symbolic capacities, and the lack of accumulation of products across generations apart from the isolated formation of traditions.

The neural systems subserving these aspects of behavior make up the main part of the brains of animals, and still make up substantial parts of the human brain. The human body possess approximately 125 degrees of freedom, our field of view is approximately 180 degrees in all directions ("panoramic vision") and our perception of the world is not only based on vision, but on complex sensory syntheses drawing on all sensory modalities (vision, audition, taste, smell, the haptic sense, and proprioception). When it is considered, that in order to support locomotion and object manipulation all this information must be processed in real time, it is evident, that the brain must have an immense processing capacity in order to accomplish these "simple" forms of behavior.

In relation to the theme of this paper, a very significant conclusion may be drawn on this basis: To the extent that the interaction can be delegated to the basic sensory-motor level for control of locomotion and object handling, immense amounts of processing power is available to the user's cognitive system, and largely without the necessity of the user's conscious effort.

Human activity mediated by material tools and linguistic signs

Humans share many locomotor and manipulative capacities with other hominids, but because of a few crucial biological differences our way of living has achieved a fundamentally different character. The neural basis for these crucial differences is too complex to review, and the task of doing so exceeds the limits of this paper. But their result is a capacity for living in complex social

Klaus B. Bærentsen

43

groups whose members can communicate verbally, construct complex objects to be used as tools for the production of specialized objects for consumption, shelter and decoration etc. On top of this emerges the ability to become conscious of one self, and reflect on ones own life activity.

In order to enter the species-specific *historically developing mode of existence*, the most significant capacities are the ability to objectify insight and knowledge in complex objects, and the ability to appropriate these objects as extensions of the body. *Artifacts* embody operational forms of past human activities, and recognized natural laws (Bærentsen 1989). The two-sided process of objectification and appropriation opens the road into history. The objectification-appropriation cycle constitutes a significant ingredient in the context of the abilities to maintain these objects and disseminate their use in the social community, communicate about them, and transmit them from one generation to the next.

The transmission of objects, symbols and competencies from one generation to the next, and in general the dissemination of acquired competencies among the members of society is achieved by learning processes in the context of social cooperation. The learning individual acts in a "zone of proximal development" under the supervision and guidance of more knowledgeable peers, who structure the actions in such a way, that the goal is accomplished, and the necessary operational capacities are developed in the individual (Vygotsky 1980).

Especially relevant in this context is the fact that the human brain is primarily adapted to control of non-symbolic locomotor and manipulative object activity (cf. Arbib 1989; Bernstein 1996) on the basis of direct perception of the layout of the physical environment and the affordances of objects (Gibson 1966, 1986). On top of this appeared the immensely important abilities for object manipulation and *constructive creation* that enabled the human species to modify nature and create the cultural-historical world including language in the modern sense (Klix 1993; Leontyev 1981, Leroi-Gourhan 1980).

The evolution of language as well as the manipulative and constructive abilities seem to be supported by conceptual understanding and representation of object characteristics (Klix 1983) that still today constitutes the functional basis of our linguistic conceptual abilities (cf. Barsalou 1998; Nelson 1974). Although adult humans have fairly sophisticated capabilities for the use of linguistic conceptual thinking in the control of their actions, the basic sense of the underlying substantial nature of objects seems to be an indispensable feature of non-pathological categorical thinking (cf. Mogensen 1997; Xu & Carey 1996). The significance and reality of this was demonstrated in the research mentioned above.

Once language had evolved, it became a powerful tool for transmission of knowledge. Modern humans grow up in cultures utilizing language for communicating and storing information and knowledge. Language makes it possible to point at, and reactivate, relevant parts of experience, that are not consciously represented in the moment, or available in the context of activity (Larsen 1989, Pulvermüller 1999).

The appropriation of artifacts in activity consists in the development of cognitive and sensory-motor capacities functioning in the same way as those developed during phylogeny. By use of linguistic signs in communication when appropriate, these functions can be activated and utilized in a productive way and by this allow tentative actions in unknown circumstances. These tentative actions may then initiate the development of skilled actions.

Summary of evolutionary arguments and some consequences

The evolutionary arguments put forward here are summarized in table 3. Briefly stated, the user's understanding of and interaction with programmed devices having screen-based interfaces are supported by "mental models".

Table 3: Aspects of human activity rooted in or acquired in different phases of the biological evolution. The first row indicates the extent of species with which humans share the kinds of activity and the objects and objectives of these activities.

EVOLUTIONARY ORIGIN	TYPICAL ACTIVITY	OBJECT OF ACTIVITY
Animalia (all living animals)	Motivated goal-directed bodily activity	Needed objects, need-related affordances of the environment
Terrestrial vertebrate (all land living animals with a vertebra)	Locomotion in the four-dimensional physical world incl. dynamic localization of objects	Life habitat, physical areas on the ground, "scenes" with affordances, obstacles, objects and other features relevant to the control of locomotion and object related activities
	Identification of objects	Recognition of objects as instances of certain natural categories and with certain operational affordances
Hominid (the highest apes)	Identification of instrumental object functions	Recognition of object functions as instrumental to the achievement certain aims (instrumental affordances)
	Manipulation and modification of objects	Natural objects with physical-functional characteristics related to needs and instrumental affordances
Homo sapiens sapiens (all humans) - as Zoon technikon (tool using creature)	Objectification- appropriation of practical-bodily cultural products(	Creation and appropriation of non-natural i.e. cultural) objects and activities and the corresponding functional systems in the brain. Concepts for abstract functions
- as Zoon politikon (social creature)	Practical and linguistic cooperation and communication with con-specifics	Coordination of own behavior in relation to activities of other members of the social group and society at large, linguistic communication (give and receive help and instructions, utilization of metaphors for explanation) Learning in social settings - appropriation of cultural object-related and social activities in social contexts. Systems of linguistic concepts
- as Zoon epi-gignoskon (self conscious creature)	Conscious reflection and meta-cognition	Perception of own activities, experiences, feelings, thoughts, cognitive activities and representations, communication and self- reflection.

These mental representations draw on cognitive and sensory-motor systems that serve to control practical activities in the four dimensional physical world (space-time). These systems evolved as adaptations to the habitat of the species during phylogenesis. Because they are adaptations, they rely on and obey some very fundamental characteristics of physical phenomena. In cognitive terms, all humans intuitively understand these characteristics of the world. Linguistic communication serves to activate available relevant sensory-motor experience, and this reactivation embody that which is normally termed "understanding".

In relation to development of interfaces, the hypothesis developed here is that much of the interaction with technology should be delegated to the level of non- symbolic processes of direct perception-action control drawing on evolutionary old "hard-wired" systems in the brain. These systems possess an immense processing power compared to symbolic information processing, and they are specially tailored for automatic control of many aspects of interaction that in current interfaces requires symbolic information processing and conscious efforts of the user.

In relation to development of interfaces, the hypothesis developed here is that much of the interaction with technology should be delegated to the level of non- symbolic processes of direct perception-action control drawing on evolutionary old "hard-wired" systems in the brain. These systems possess an immense processing power compared to symbolic information processing, and they are specially tailored for automatic control of many aspects of interaction that in current interfaces requires symbolic information processing and conscious efforts of the user.

Perceptual information for intuitive interfaces

According to Gibson (1966, 1986) the visual perception of the physical world is based on the *information* carried by *the ambient optic array*. The ambient optic array is a term that denotes the total amount of light that is entering the eyes of an observer. This light is structured by the innumerable reflections it has passed through on its way from the sun, via the atmosphere onto the ground, and the diverse surfaces of objects it may have encountered and been reflected by, before it reaches the eye of an observer. On every encounter with molecules in the air or on the surfaces of objects, the light is changed (i.e. structured) either with regard to its direction, its intensity (brightness) or its spectral composition (color). The changes are determined unambiguously by the structure and texture of the reflecting surfaces, and the light that enter the eyes of an observer is thus carrying information about, or specifies, the nature of the reflecting surfaces.
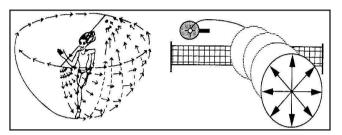
The structural changes that are caused by the reflecting surfaces, depends on their angle in relation to the direction of the incoming light, as well as on their texture. Delimited surfaces are thus specified by delimited sections of the total optic array and carry the "stamp" from the texture of the specific reflecting surface.

When we move in the physical world, we have a conscious notion of the goal towards which we move or strive. But the operational details of the activity are determined by the perceptual pickup of available information about the conditions for the activity. This information is carried by the transformation of the optic array according to the laws of "ecological optics", as described by Gibson (1966, 1986), Gibson, Olum & Rosenblatt (1955), Lee (1980), Turvey & Carello (1986) and others. Some crucial aspects of these laws are summarized in the following paragraphs.

Events are specified by patterns of change in the optic array, as illustrated in (Figure 14), showing
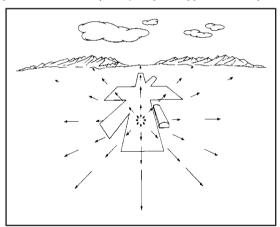
two prototypical situations: an actor moving around in the world, and an object moving towards the observer. Arrows illustrate the displacement of elements in the optical array constituting the field of vision during a certain amount of time.

Figure 14: Global and local transformations in the field of vision corresponding to subject locomotion and movement of objects (Turvey 1990 p. 942).
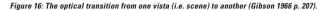

When the subject moves through the environment, the optical elements in the whole field of vision are expanding from the point towards which he moves, passing by the edges in the periphery of his field of vision. Behind the subject the optical flow field is converging towards the point directly opposite of the heading. When an object moves towards the subject, the local field of vision corresponding to the object is expanding, as seen to the right in the figure.
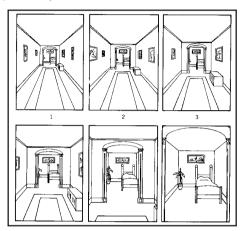
Figure 15: The outflow of the optic array during a landing glide (Gibson 1986 p. 125).

The optic flow during locomotion can be represented by the translation vectors of elements in the visual field of an observer during a specified amount of time. This is exemplified in (Figure 15) for a situation in which a pilot is approaching the runway during landing of an aircraft. As it is seen in the figure, the optic array is flowing out from a point corresponding to the heading of the aircraft. The optical flow may be described as a progressive global magnification of optical texture from the focal point towards the periphery of the visual field. The distances of the objects from the observer, their angular distance from the center of the visual field, and the velocity of the observer

Klaus B. Bærentsen

47

towards the point of landing determine the magnitude of the flow-vectors in different parts of the visual field. What is very important here is the circumstance, that the objective relations between the observer and the environment in terms of the relative distances and velocities are unambiguously specified in the optic flow.

Figure 16: The optical transition from one vista (i.e. scene) to another (Gibson 1966 p. 207).
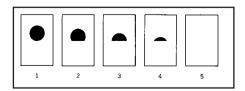

During locomotion the transformation of the optical array is determined by the layout of the environment and the characteristics of the observers loco motion. This is shown in the depiction of the transformations caused by the movement of an observer from one vista to another (Figure 16). In this example the doorframe delimits an opening disclosing a scene behind the wall. During the approach to the opening the edges of the frame gradually disclose or "disoccludes" more and more details of the room behind. Such occluding and disoccluding edges serve a very important function for perception during navigation in the environment, as they inform about the relative position of objects in 3D space.

When an object moves in the field of vision relative to a static observer and the surroundings, the transformations in the optical array are local, and only cover a part of the visual field corresponding to the visual angle in the field of view of the observer, whereas the global field of vision remains constant. If the object is moving towards the observer, the changes will consist in a progressive expansion of the part of the visual field corresponding to the object.

Figure 17: Five "snapshots" of the local transformations of the optic array specifying the disappearance of an object behind an edge (Gibson 1966 p. 204).

48

Klaus B. Bærentsen

This will cause a progressive deletion of texture from the background, at the edges of the object. If the object is moving at a constant distance from the observer the changes can be described as a progressive deletion or occlusion of optical texture from the background at the leading edge of the object, and a progressive disocclusion at the trailing edge. When an object passes behind another, the defining texture disappears at the edge of the occluding object as shown in (Figure 17).

If the user's navigation in the functional space of a system is to be supported by the capacities for direct perception, the transformation of the information on the interface as he "moves" around in the functional space of the system (i.e. from one menu page to another) must be in accordance with the laws of "ecological optics". A global or local translation of texture in accordance with the "laws" of ecological optics unambiguously specifies a certain event in a "virtual space", and will be perceived as such. This has been documented for a number of specific manipulations of the visual input to experimental subjects (Gibson 1986, ch. 9-11; Gibson et al. 1955; Lee 1980), and this effect has found large-scale application in for instance flight simulators, and other kinds of simulators, as well as in arcade and computer games.

IMPLEMENTING INTUITIVE INTERFACES

In order to construct user interfaces that support the utilization of the general human cognitive and sensory-motor competencies for locomotion and object manipulation, it is necessary to analyze the various functions in the system for which the interface is constructed, and assign them to the categories of scenes, objects and actors, characterized by basic properties in relation to interaction.

Objects and scenes are the most crucial part of the task, and because they are most easily implemented, they will be treated first. These functions will only require powerful graphical processing power in the equipment. The last topic - functions that may be characterized as actors - will not be treated at any length in this paper. Today only simple aspects of these functions are implemented in existing AV products. Although the presentation of instructions and feedback information is often described as "communication", the capacity of contemporary technology is far from the level of sophistication needed to support "intuitive" interaction as conversation (Larsen 1989; Shneiderman and Maes 1997; Suchman 1987).

Changing the interface of the Avant TV

As a concrete example a preliminary analysis of some aspects of the interface for a standard B&O Avant TV will be presented here. The Avant is a rather complex integrated AV system with TV and programmable VCR (Video Recorder), and optional Satellite, Dolby surround sound, PIP ("Picture In Picture"), and various other functions. It may be programmed to turn to two alternative viewing angles automatically, and is capable of being the central unit in a B&O link system.

In table 4 the system of hierarchically organized menus and sub-menus on the interface of a standard Avant TV is listed, and the kind of functions on each menu page today is indicated in square brackets. The words used to denote the type of functions, and their approximate translation in the preliminary analysis is as follows: **Choice** denotes the possibility to go to a submenu, and would translate to an "opening" or a "door".

Klaus B. Bærentsen

49

Table 4: The system of Menu's on the B&O Avant TV

MENU HIERARCHY	FUNCTION	METAPHORICAL CONTENTS
Main menu	[choice]	Scene, 5 openings
1. Timer record	[choice]	Scene, 2 openings, object
1.1 Timer index (recordings)	[info + fill-in]	Scene, 1 opening, object
[1.2] Teletext Programming	[info + fill-in]	Scene, object
2 Timer play	[fill-in + choice]	Scene, 2 openings, object
2.1 Timer index (play/standby)	[info + fill-in]	Scene, 1 opening, object
[2.2] Teletext Programming	[info + fill-in]	Scene, object
3 TV list (list of channels)	[info + fill-in]	Scene, n openings, object
4 Sat list (list of channels)	[info + fill-in]	Scene, m openings, object
5 Setup	[choice]	Scene, 8 openings
5.1 Tuning	[choice]	Scene, 3 openings
5.1.1 TV (tuning)	[fill-in + choice]	Scene, 1 opening, object
5.1.2.1 Fine tune	[fill-in]	Scene, object
5.1.2 Sat (tuning)	[fill-in + choice]	Scene, 1 opening, object
5.1.2.1 Fine tune	[fill-in]	Scene, object
5.2 Sound	[choice]	Scene, 2 openings
5.2.1 Adjust	[fill-in]	Scene, object
5.2.2 Speakers	[fill-in]	Scene, object
5.3 Picture	[fill-in]	Scene, object
5.4 Stand	[fill-in]	Scene, object
5.5 Source	[fill-in]	Scene, object
5.6 Menu	[fill-in]	Scene, object
5.7 Clock	[fill-in]	Scene, object
5.8 V.Tape	[choice]	Scene, 2 openings
5.8.1 Adjust	[fill-in]	Scene, object
5.8.2 Basic Setup	[fill-in]	Scene, object

The number of openings corresponds to the number of accessible sub-menus. For the channel lists the number of openings is indeterminate, as it depends on the number of channels that are tuned in on the set. **Fill-in** indicates a function which the user can adjust or program by entering data in specified fields, it translates to "object that can be manipulated", or "actor with memory to interact or communicate with" depending on the exact type of the function. **Info** indicates a menu where the user can access some information; it would translate to a view of static "objects", or simple statements from "actors".

Examples of the kinds of items on the user interface that may be represented by the various metaphors are tentatively summarized in table 5.
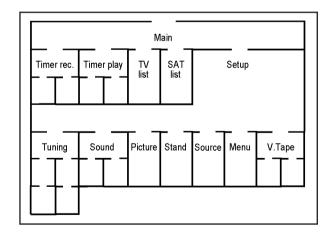
If the menu system of the Avant TV were actually analyzed this way, the structure would change because many transitions between scenes would become superfluous, as some scenes could contain more objects. The resulting change of architecture is not taken account of here however. For our purpose the architecture will be held constant.

Table 5: Interpretation and categorization of items on screen menus in terms of objects, scenes and actors.

METAPHOR	ITEM ON INTERFACE	EXAMPLES
Object	Passive functions on fill-in menus	TV- and SAT-Channels on the Channel lists
-	- concrete	Clock
-	- functional/relational	Synchronization of clock to Teletext
Scenes	Menu pages as the background on which functions are presented	Main menu, Setup menu, Tuning menu, TV (tuning) menu, Channel List etc.
- openings	Labels for sub-menus on choice menus	"Tuning" on the Setup menu
Actors	Active functions (e.g. programming) on fill-in menus	Timer programming of VCR recording Timer play programming
- information -	simple instructions and feedback information	"Synchronize clock to Teletext?" "Press MENU for teletext Programming" "Programming stored"
- dialogue	Interactive help functions	n.a.

The hierarchically organized system of menus may be transformed to a corresponding system of scenes and openings as illustrated in Figure 18. This may make it easier to imagine how transitions from one menu page to another can be illustrated as moving from one room (scene) to another.

Figure 18: The system of menus transformed into an organized set of scenes.

Some features of the representations of different categories of items on the interface are summarized in table 6 and table 7. These features are derived from the knowledge about our understanding and perception of objects and situations as it has been described in the previous sections.

Klaus B. Bærentsen

51

Table 6: Functional features of items on interfaces interpreted as objects, scenes and actors

METAPHOR	ITEM ON INTERFACE	CHARACTERISTIC FEATURES
Object	Passive functions on fill-in menus	2D projections (i.e. "textured surfaces") of "substantial" objects preserving their identity during manipulation and relocation. Projections must obey certain principles: 1) *Cohesion*. Surfaces belong to a single object, if and only if they are attached to each other. If they are not attached to each other they do not belong to a single object. Objects are circum scribble as continuous surfaces. 2) *Contact:* Surfaces move together, if and only if they are in mutual contact. Surfaces that are not in mutual contact do not move together. 3) *Continuity*. Movements of objects are continuous in time and space. Separate objects move in separate trajectories that are continuous in time and space. Two objects can't be at the same place at the same time.
-	-concrete	"Substantial" objects, with perceptible affordances that can be illustrated (e.g. technical elements like buttons to press, rotate, slide).
-	- functional	"Abstract" objects with imperceptible affordances that cannot be illustrated (e.g. LNB, link-frequency)
Scenes	Menu pages as the background on which functions are presented	2D perspective projections of rooms or scenes by appropriately textured surfaces. When moving between menus the visual representation on the screen is transformed in a continuous manner as a geometrical projection of the underlying 3D architecture of the system. The transformation informs the user about his path through the system.
Openings	Sub-menus on choice menus	"Openings" are local textures corresponding to the projection of other "scenes" (i.e. sub-menus) "behind" the pres ent, and thus allow a preview to their contents
Actors - executive	Active functions (e.g. programming) on fill-in menus	Objects accepting instructions and carrying out corresponding functions at specified times. Must allow flexible entering of data, in a format that is congru ent with the users point of view. Setting of parameters should be supported by appropriately chosen defaults. Must provide appropriate status information etc.
- information	Instructions	The information necessary to present the available affordances and registered programmings that are not directly perceivable.
- dialogue	Help and guidance	Dialogue- reflection- and learning systems to support learning in the zone of proximal development. Natural language dialogue systems needs a capacity for "repair of problems of understanding".

Objects would be represented as local and circumscribed textured parts of the interface, scenes as global (i.e. full screen) texture providing the background for projections of objects. Dynamic situations, i.e. state transitions would be represented as local transformations (objects) vs. global transformations of textures on the interface (for scenes).

When we move in known terrain and operate devices with known functions we can mostly depend on perceptually available information. But when we get into unfamiliar conditions or operate unknown complex devices, we may need an explanation, e.g., a map or an instruction. It may be necessary to train our perceptions in order to evoke adequate mental models' schemas and maps of the world (cf. Neisser 1976). The use of text and textual information is obviously necessary when it concern functions, that don't have a physical form that can be illustrated with a picture or reveal itself during dynamic transformations of the information displayed on the interface as it is envisioned here.

A properly formulated text may work as a set of "fishing hooks" into the memory of the user and retrieve relevant information from the prior experience. This information may make it possible by analogy to build a mental model containing elements that make it possible to get an understanding of the current situation and the available appropriate actions (c.f. Bærentsen 1996, Gentner & Holyoak 1997, Gentner & Markman 1997, Holyoak & Thagard 1997, Larsen 1989).

Table 7: Immediately perceptible "ecological information" specifying objects and scenes on screen based interfaces

METAPHOR	CHARACTERISTICS ON INTERFACE
Objects (as static entities)	Objects are visible as textured surfaces that together create a form of a certain relative size. An object in front of another object occludes the other object in part or completely
Objects (as dynamic entities)	Attributes may be changed, but they "stick together" if the object is moved. The object maintains its "substantial" identity over modifications. When it is moved, the configuration of the textured surface defines a continuous local "flow field" determined by the 3D structure of the object, the course of its movement relative to the background scene and other objects. The texture of the background will be occluded at the leading edge of the object, and disoccluded at the trailing edge. The flow field can be described by a set of local expansion, translation and rotation vectors for the corresponding local textured surface.
Scenes (as static entities)	A scene (vista) is a perspective view of a 3D spatial scenery. The user is at the point of observation defining a "point of disappearance" towards which parallel lines "into" the scene converge according to the normal laws for perspective representations. The surfaces defining the scene are textured; the texture is graded toward the point of disappearance.
Scenes (as dynamic entities)	When the users point of observation is changed, the course and speed of the movement defines a continuous and global flow field determined by the 3D structure of the scene and any object coming into the field of vision. The flow field can be described with a set of expansion- translation- and rotation vectors
Openings	Surfaces in scenes may have openings leading to other vistas. Openings are local textures corresponding to at perspective projection of surfaces in the scene to which the opening leads. During transitions between the scenes, the local texture gradually expands and occludes more and more of the surrounding texture, revealing an increasing amount of texture corresponding to the new scene.

Klaus B. Bærentsen

53

SUMMARY AND REMAINING QUESTIONS

The main argument in this paper is that many current technological systems with menu system interfaces present symbolic information that require reading of textual status and feedback information, and intellectually guided setting of abstract parameters for symbolized functions. Many aspects of the interaction with such systems would be better served by graphical displays presenting dynamic information in accordance with the principles elucidated in ecological optics (Gibson 1986; Gibson, Olum & Rosenblatt 1955; Lee 1980; Turvey & Carello 1986). This would allow the user to draw upon the massive processing powers of the perception-action systems.

Information about "scenes" and "objects" can be presented on screen-based interfaces by following what is known about the principles of ecological optics. By utilizing these principles it is possible to inform the user about movements around in the functional space of the system and about handling objects as "substantial entities" that can be displaced and manipulated.

It is possible to display the affordances of "scenes" as explorable rooms or places in a space, of walls as obstacles, openings in walls as passages to other scenes. The dynamic transformations are a major source of information about the static 3D layout of the "environment" (i.e. the functional space) and the relative position of objects and the observer in this environment. But what are the limitations due to the fact that the field of view is not panoramic when the scene is displayed on a screen of limited size? And what does it mean that the observer doesn't move, so there is only visual exproprioceptive information available about the movement in space?

Information can be displayed about objects as "graspable", "displaceable" and "manipulable" or the opposite. But what about the perceivability of technical affordances (or functional as different from the "substantial" affordances)? Technical affordances may be directly perceptible if they are associated with the surface characteristics of the object, or if the user has acquired the relevant knowledge about the functions of the object, and how they are signaled by specific perceptible surface characteristics. There is need for a kind of "optics of technical affordances" or "technological optics" (as a concomitant to the "ecological optics") describing the directly perceivable characteristics that signal standard affordances to anybody who has grown up in a technological environment. Some examples are buttons that are pushable, turnable or slideable, chairs that are "sitable", wheels that are "rollable", icons that are recognizable and texts that are readable.

But the argument is not that every aspect of the user interface should be delegated to a perceptual level. Some functions are better served by a linguistic instruction or an explanation of conceptual affordances. It also remains to be determined what minimal knowledge about technological affordances can be supposed to be generally available to the "normal users"? What are the relevant basic level categories (cf. Rosch 1978) on which it is possible to base illustrations and linguistic explanations? How are these basic level categories determined by the pragmatic context of their development and retrieval (cf. Barsalou 1993)? And what are the relevant models of knowledge on which to build more advanced help and explanation functions in the systems (cf. Hacker & Jilge 1993).

ACKNOWLEDGEMENTS

Notes

[1] Bang & Olufsen is a Danish manufacturer of High-end Audio-Video equipment.

References

Andersen, P. B. & K. H. Madsen (1989): Design and professional languages. In: Andersen, P. B. & T. Bratteteig (eds.): *Computers and Language at Work.* Institute of Informatics Research Reports no. 126, Århus, pp. 157-196.

Arbib, M. (1989): *The Metaphorical Brain 2.* Wiley, N.Y. etc.

Bannon. L. (1986): Helping Users Help Each Other. In: Norman, D. A. & S. W. Draper (eds.): *User Centered System Design*. Hillsdale: Erlbaum, pp. 399-410.

Bardram, J. (1997): Plans as Situated Action: An Activity Theory Approach to Workflow Systems. In: *Proceedings of ECSW'97*, Lancaster.

Barsalou, L. (1993): Flexibility, Structure, and Linguistic Vagary in Concepts: Manifestations of a Compositional System of Perceptual Symbols. In: Collins, A. F. et al. (eds.): *Theories of Memory. Erlbaum*, Hove etc. pp. 29-101.

Barsalou, L. (1999): Perceptual Symbol Systems. *Behavioral and Brain Sciences*, Vol. 22, pp. 577-660.

Bedny, G. & D. Meister (1997): *The Russian Theory of Activity. Current Applications to Design and Learning*. L. Erlbaum Mahwah, N.J. etc.

Bernstein, N. A. (1996): On Dexterity and its development. In: Latash, M. L. & M. T. Turvey (eds.): *Dexterity and its Development*. Erlbaum, Mahwah, New Jersey, pp. 1-244.

Brüel, S. & N. Å. Nielsen (1987): *Gyldendals Fremmedordbog*. Gyldendal, København.

Bærentsen, K. B. (1989): Mennesker og maskiner. In: Hedegaard, M.; V. R. Hansen & S. Thyssen (red.): *Et virksomt liv. Udforskning af virksomhedsteoriens praksis*. Aarhus Universitetsfor-lag, Århus. pp. 142-187.

Bærentsen, K. B. (1992): Sikkerhedsmæssige konsekvenser af informationsteknologien. In: Bærentsen, K. B.; F. Øwre, E. Relster & S. Munck: *Informationsteknologi. Informationsteknologiens menneskelige konsekvenser for fremtidens virksomhed. Konference mandag d. 4. maj 1992. Danmarks tekniske Højskole*. Dansk Automationsselskab, Lyngby, pp. 4-32.

Bærentsen, K. B. (1996): Episodic Knowledge in System Control. In: Andersen, P. B.; B. Holmqvist, H. Klein & R. Posner (eds.): Signs of Work: *Semiosis and information processing in organisations*. Walter de Gruyter & Co. pp. 283-323.

Bærentsen, K. B.; J. Kvorning & L. Skov (1989): Verbal Reports on Control Actions in Power Plants. In: *Proceedings of the Eight European Annual Conference on Human Decision Making and Manual Control. DTH, Lyngby Denmark, June 12-14, 1989*. pp. 292-314. Technical University of Denmark, Lyngby.

Bærentsen, K. B. et al. (1995): *Undersøgelse af almindelige brugeres anvendelse og forståelse af videorecorderes funktioner*. PIAU, Århus (unpublished manuscript in Danish).

Bærentsen, K. B. & H. Slavensky (1999): Usability - A Contribution to the Design Process. *Communications of the ACM*. Vol. **42** no. 5, pp. 73-77.

Bødker, S. (1987): *Through the Interface - a Human Activity Approach to User Interface Design*. DAIMI PB-224, Aarhus.

Carroll, J. M. &R. L. Mack (1985): Metaphor, computing systems, and active learning. *International Journal of Human-Computer Studies* Vol. **22**, pp. 39-57. (Reprinted in: International Journal of Human-Computer Studies Vol. 51, 1999, pp. 385-403.)

Carroll, J. M. & J. C. Thomas (1982): Metaphor and the Cognitive Representation of Computing Systems. *IEEE Transactions on Systems, Man, and Cybernetics*. Vol. **SMC-12** (2), pp. 107-116.

Chafe, W. (1990): Some Things That Narratives Tell Us About the Mind. In: Britton, B. K. & A. D. Pellegrini (eds.): *Narrative Thought and Narrative Language*. Erlbaum, Hillsdale, N.J. pp. 79-98.

Chan, T.-C. & R. E. Shaw (1996): What is Ecological Psychology? *Psychologia*, Vol. 39, pp. 1-16.

Damasio, A. R. (1989): Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition.* Vol. 33, pp. 25-62.

Deecke, L. (1996) Planning, preparation, execution, and imagery of volitional action, (Introduction / Editorial) in: L. Deecke W. Lang, A. Berthoz (Eds.) Mental Representations of Motor Acts. *Cognitive Brain Res 3* / Special Issue (2) 59-64.

Donald, M. (1991): *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge University Press, Harvard

Fabri, K. E. (1983) [A.N. Leont'evs scientific heritage and questions of the evolution of the psyche. - in Russian]. Fabri, K. É.: Naučnoe nasledie A. N. Leont'eva i voprosy èvoljucii psichiki. V sb.: *A. N. Leont'ev i sovremennaja psichologija*. (Pod red. A. V. Zaporožca, V. P. Zinčenko, O. V.Ovčinikovoj, O. K. Tichomirova). MGU. Moskva, 1983, s. 101-117.

Farah M (1984): The neurological basis of mental imagery: A componential analysis. *Cognition*, Vol. **18**, pp. 245-272.

Gentner, D. & K. J. Holyoak (1997): Reasoning and Learning by Analogy. *American Psychologist*. Vol. **52**(1), pp. 32-34.

Gentner, D. & A. B. Markman (1997): Structure mapping in analogy and similarity. *American Psychologist*. Vol. **52**(1), pp. 45-56.

Gibson, J. J. (1966): *The Senses Considered as Perceptual Systems*. Houghton Mifflin, Boston.

Gibson, J. J. (1986): *The Ecological Approach to Visual Perception*. Lawrence Erlbaum, Hillsdale.

Gibson, J. J.; P. Olum & F. Rosenblatt (1955): Parallax and Perspective During Aircraft Landings. *American Journal of Psychology*. Vol. **68**, pp. 372-385.

Gregory, R. L. (ed.) (1989): *The Oxford Companion to the Mind*. Oxford University Press, Oxford.

Hacker, W. & S. Jilge (1993): Vergleich vershiedener Methoden zur Ermittlung von Handlungswissen. *Zeitschrift für Arbeits- u Organisationspsychologie*. Vol. **37**(N.F.11)2, pp. 64 - 72.

Holyoak, K. J. & P. Thagard (1997): The analogical mind. *American Psychologist*. Vol. **52**(1), pp. 35 -44.

Jeannerod, M. (1994): The Representing Brain: Neural Correlates of Motor Intention and imagery. *Behavioral and Brain Sciences*, Vol. **17**(2), pp. 187-245.

Johnson, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: Chicago University Press.

Klix, F. (1993): *Erwachendes Denken. Geistige Leistungen aus evolutionspsychologischer Sicht*. Spekturm Akademischer Verlag, Heidelberg etc.

Larsen, S. F. (1989): Communication of Information as Contextualised Activity. In: Finneman, N. O. (ed.): *Proceedings from the symposium "Theories and technologies of the knowledge society"*. Center for Cultural Research, Aarhus, September 21, pp. 11-32.

Lee, D. (1980): The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society of London, B 290*, pp. 169-179.

Leontyev, A. N. (1977): Activity and Consciousness. In: *Philosophy in the USSR. Problems of Dialectical Materialism*. Progress, Moscow, pp. 180-202.

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

Leont'ev, A. N. (1979): The Problem of Activity in Psychology. In: Wertsch, J. V.: *The Concept of Activity in Soviet Psychology*. Sharpe, N.Y. pp. 37-71.

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

Leontyev, A. N.; A. R. Luriya & A. A. Smirnov (eds.) (1966a): *Psychological Research in the U.S.S.R.* Progress, Moscow.

Leontyev, A. N.; V. P. Zinchenko, & D. Yu. Panov (eds.) (1966b): *Engineering Psychology*. Air Force Systems Command. Foreign Technology Division.WP-AFB, Ohio.

Lomov, B. F. (1963): *Man and Technology: Studies of engineering psychology*. U.S. Dept. of Commerce, Office of Technical Services, Joint Publication Research Service, Washington D.C.,

Mammen, J. (1989): The relationship between subject and object from the perspective of Activity Theory. In: N. Engelsted, L. Hem & J. Mammen (eds.): *Essays in general psychology. Seven Danish contributions*. Århus: Aarhus University Press, pp. 71-94.

Mammen, J. (1993): The elements of psychology. I: N. Engelsted, M. Hedegaard, B. Karpatschof & A. Mortensen (eds.): *The societal subject*. Århus: Aarhus University Press, pp. 29-44.

Mammen, J. (1994): Rubinstein's conception of the "leap" to the specifically human consciousness in Sein und Bewusstsein: A critical evaluation. *Multidisciplinary Newsletter for Activity Theory*, 15/16, pp. 29-32.

Klaus B. Bærentsen

Milner, A. D. & M. A. Goodale (1998): Precis of: A. D. Milner & M. A. Goodale The Visual Brain in Action, *Psyche,* Vol. **4**(12), October http://psyche.cs.monash.edu.au/v4/psyche-4-12-milner.html

Mogensen, J. (1997): Spædbarnet og den skizofrene. En verden til forskel. *Agrippa*, Vol. **18**, pp. 22-35.

Myers, B. A. (1998): A Brief History of Human-Computer Interaction Technology. *Interactions*, March+April, pp. 44-54.

Neisser, U. (1976): *Cognition and Reality*. Freeman, San Francisco.

Neisser, U. (1994): Multiple Systems. *European Journal of Cognitive Psychology*. Vol. **6**(3), pp. 225-241.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. *Psychological Review*, Vol.**91**, pp. 267-283.

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

Oxford (1989): *Oxford Advanced Learner's Dictionary* (New Edition), Oxford University Press, Oxford.

Perkins, R.; D. Smith Keller & F. Ludolph (1997): Inventing the Lisa User Interface. *Interactions*, January+February, pp. 41-53.

Pulvermüller, F. (1999): Words in the brain's language. *Behavioral and Brain Sciences*. Vol. **22**, pp. 253-336.

Reed, E. S. (1996): *Encountering the World. Towards an Ecological Psychology*. Oxford University Press, New York etc.

Robertson, B. (1999): Biz viz gets real. *Computer Graphics World*. April, pp. 29-34.

Rosch, E. (1978): Principles of categorization. In Rosch, E., and Lloyd, B. (eds.) *Cognition and Categorization*. Hillsdale, N.J.: Erlbaum, pp. 27-40.

Rubinshteijn, S. L. (1976): [Psychological problems in the works of Karl Marx. -in Russian. Original published 1934]. Rubinštejn, S. L.: Problemy psichologii v trudach Karla Marksa. V Rubinštejn, S. L.: *Problemy obščej psichologii*, Pedagogika, Moskva s. 19-46.

Schmandt-Besserat, D. (1978): The Earliest Precursor of Writing. *Scientific American*. Vol. **238**(6), pp. 50-59.

Shneiderman, B. (1983): Direct Manipulation: A Step Beyond Programming Languages. *IEEE Computer*, August, pp. 57-69.

Shneiderman, B. & P. Maes (1997): Direct Manipulation vs. Interface Agents. *Interactions*. November + December, pp. 42-61.

Suchman, L. A. (1987): *Plans and Situated Actions. The problem of human-machine communication*. Cambridge University Press, Cambridge etc.

Turvey, M. T (1990): Coordination. *American Psychologist*. Vol. **45**(8), pp. 938-953.

Turvey M. T. & C. Carello (1986): The Ecological Approach to Perceiving-acting: A Pictorial Essay. *Acta Psychologica*, Vol. **63**, pp. 133-155.

Velichkovsky, B. M. (1994): The Levels Endeavour in Psychology and Cognitive Science. In: Bertelson, P.; P. Eelen & G. d'Ydewalle (eds.): *International Perspectives on Psychological Science. Vol. 1: Leading Themes*. Erlbaum, Hove etc. p. 143-158.

58   Klaus B. Bærentsen

Vygotsky, L. S. (1966): The development of the higher mental functions. In: Leontyev, A. N.; A. R. Luriya & A. A. Smirnov (eds.): *Psychological Research in the U.S.S.R*. Progress, Moscow, pp. 11-45.

Vygotsky, L. (1980): *Mind In Society*; Cambridge University Press, Harvard

Wharton, C.; J. Rieman; C. Lewis & P. Polson (1994): The Cognitive Walkthrough Method: A Practitioner's Guide. In: Nielsen J. & R. L. Mack (ed.): *Usability Inspection Methods*. Wiley, New York etc. pp. 105-140.

Xu, F. & S. Carey (1996): Infants' Metaphysics: The Case of Numerical Identity. *Cognitive Psychology*. Vol. **30**, pp. 111-153.

Zinchenko, V. P. & V. M. Munipov (1989): *Fundamentals of Ergonomics*. Progress, Moscow.

60