Presentation media for product interaction

Piet Westendorp

Presentation media for product interaction

Presentation media for product interaction

Proefschrift
ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus, prof. dr. ir. J.T. Fokkema,
voorzitter van het College van Promoties,
in het openbaar te verdedigen op
maandag 10 juni 2002 om 16:00 uur
door

Pieter Hendrikus WESTENDORP
doctorandus in de Nederlandse Taal- en Letterkunde
geboren te Baarn

Dit proefschrift is goedgekeurd door de promotoren:
Prof. P.P. Mijksenaar
Prof. dr. J.M. Dirken Eur Ing.

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. P.P. Mijksenaar, Technische Universiteit Delft, promotor
Prof. dr. J.M. Dirken Eur Ing., Technische Universiteit Delft, promotor
Prof. dr. J. Aasman, Technische Universiteit Delft
Prof. dr. D.G. Bouwhuis, Technische Universiteit Eindhoven
Prof. dr. J.W. Drukker, Technische Universiteit Delft
Prof. dr. C. Strothotte, Hochschule Magdeburg, Duitsland
Dr. K.M. van der Waarde, Technische Universiteit Delft, adviseur

If it's not easy to write about, it's not easy to use.
So I use writing as a guide for design.
Jef Raskin (Apple no. 34)

ISBN 90-9015900-2

Cover illustrations

Cataract operation. 12th Century. Manuscript Sloane. London Library (Herrlinger, 1967, p. 17).

Do not use lighters. US Air - Boeing 757 Safety card

Arrow on sewing machine. Naumann KL-14, late 19th century.

Philips TBX telephone operator manual. Flowchart.

CHAPTER 1

A motivation from practice and from theory for this study

1.1. Practice

User instructions of complex technological products have become an integrated part of daily life. Whether they are presented as printed user manuals, on-line help systems, on packages, or as traffic signs, instructions are nowadays indispensable for people to install and use products and software; open packaging; orientate themselves in countries, cities, airports and railway stations; taking medicine, filling in forms or using automated services. Present-day products are so versatile and so complex, and personal help is so limited, that extra help is needed, in spite of excellent user interfaces. These products cannot speak for themselves; assistance has become a prerequisite for the user.

These instructions have to help a wide variety of possible users of the products, surroundings, forms and services. Even for one specific product, the number of potential users may be enormous, and these users may have widely-ranging cultural backgrounds, education, knowledge, technical feeling and other aspects that make it difficult to design, write and illustrate instructions that are comprehensible for every single user. Traffic signs have to be understood by almost everybody on the planet, Microsoft Office is used by over 250 million people, every year many millions of passengers have to find their way at Kennedy International Airport, millions have to know what to do in case of an emergency in an airplane or learn to drive their new Volkswagen; and hundreds of thousands have to construct their identical Ikea chairs, open the same jam jars or take the same medicines. Users may be professionals or laymen, experienced or first-time users, young or old, highly motivated to learn or totally uninterested, or even unwilling to learn.

User instructions can have two main functions: informing and instructing. They may inform about the functionality, exceptional use and abuse, maintenance, transportation, storage, replacement and removal of the product after use. Moreover, they may contain legal and commercial information. Of course, user instructions have to instruct about safety, how to setup or install and use the product. The degree to which users will need information or instruction may vary greatly. Some will hardly need any information (or think they do not need it); others may only need some information and try to avoid the instructions. Some may need and indeed study the complete user manual; others will not have a look at it at all. Some will use the instructions only after they have tried out everything on the user interface and had to give up, others will start by carefully studying the user manual. The designer has to take all these possible users and their needs into account when designing the instructions, and make them efficiently and effectively so they are understandable and attractive for all of these users - a very difficult task.

The user instructions that come with a product should not be considered as stand- alone information. It is part of a stream of information that users may receive about the product, starting with advertising or oral communication from their neighbor to tutorials and training courses, and includes the user interface. For the designer of instructions it is difficult to guess what a priori idea the buyer of the product has about its functions and ways to install and operate it. Some users may indeed have an unexpected or even totally wrong idea about the functions of the product or the way to operate these functions. With these unexpected or even wrong ideas they start using the product, and the user manual.

Designers have produced a wide variety of user instructions. They do not only vary along with the type of product or the variation within this type (a new version every year), but also for the same types of products, designers have produced a wide variation of user manual designs. For instance, the safety instruction cards in passenger airplanes basically need to convey the same message. But it appears that they do not contain the same information (different selections are made) and we can distinguish a variety of design concepts, detailing and styles.

It is unlikely that anyone would buy a product for the pleasure of studying the user manual (except for those who collect user manuals), or that many people enjoy traveling because of the puzzle to find the right subway or correct way to use the ticket-vending machine. Consulting user instructions is considered a requisite suffering to use a new product that will make life so much nicer. And indeed, user instructions may not always consist of the nicest pictures or texts. As a result of all of this, user instructions are notorious for their incomprehensibility.

However, it is questionable whether user manuals in general really are so bad. One easily forgets that the products are very complex, sometimes have awkward user interfaces or that the designer has a completely different idea of what a good user interface should look like or what user instructions should tell or how these should be structured and designed. Moreover, user interface designers often have to deal with other constraints, such as the obligation to further develop designs on the basis of previous designs so users of previous versions will not be deterred with a completely new design. In Microsoft Windows[1], for instance, one can easily discern the DOS background and this has put constraints on the design of Windows; it would possibly have been much more user-friendly if it would have been designed without the idea that DOS users should be able to recognize the DOS commands. Likewise, designers sometimes have to take demands from the marketing department into account. Microsoft Word, for example, became more complex because of the possibility for users to make it look and feel like WordPerfect, which was obviously requested by the marketing department.

In summary, user instructions have to inform and instruct a wide variety of users about a wide variety of complex products, packaging, surroundings and services in several stages of the use of products. They are part of a complex information process and may be presented in many ways. Products have become very versatile while personal communication about them has become less. Moreover, user interfaces are not always, and maybe cannot always be, self-explaining. Therefore, users often have to rely on user instructions, and they have very different a priori ideas about both the product and the instructions. This, of course, causes problems in the communication.

1.2. Research

The widespread use, the increasing necessity to rely on user instructions, and the irritations that the use of instructions seems to

[1] Excluding NT versions.

cause, justify research into the best ways to design user instructions. Indeed many studies have been carried out, as can be seen from the overviews by Schriver (1989 and 1997). These studies concern many aspects of user instructions: the way instructions are used, the structure and formulation of texts and the design of pictures, the content, the relation with user interfaces, educational and organizational aspects, etc. The ultimate goal of all these studies is the improvement of the effectiveness, efficiency or attractiveness of user instructions - a motivation from practice.

A second reason to study user instructions is more theoretical. Studying the way user instructions function may give an interesting view on the way people interact with complex products. Human-machine interaction is often considered to be the interaction between user and user interface. User assistance (all supporting information that may help the user, including user manuals, reference cards, on-line help, guided tours, tutorials, etc.) is hardly ever taken into account, although in daily life users often need to consult some kind of user assistance, often the user manual. This means that the user does not only interact with the product user interface, but that a third element is involved: the instructions, which, by themselves may be an extra cognitive load. To describe the comprehensive human-machine interaction, a model must be developed that includes this extra element and the switching that users have to do between the user interface and the user instructions.

Moreover, a study of such a more comprehensive interaction process may shed a different light on various aspects of information processing theories, such as differences in perception of user interface and user instructions. It may also elucidate aspects concerning the cognitive process in human-machine interaction, such as problem solving, decision-making, memory and the forming of a mental model or device model. In the present study perceptive and cognitive aspects of the switching that users have to do between the spatial-graphical user interface and the linear-procedural instructions will be discussed at length. This will result, for example, in the development of a different view on mental models that users may develop of products they use. It will also result in a discussion about the relative efficiency of the direct recognition that an image offers versus the direct abstraction that text offers.

1.3. Relative efficiency of text, pictures and animations in user instructions

This study deals with the relative efficiency of various presentation media - text, still pictures and animations - for user instructions. This has been the subject of many contrastive studies, which have, however, received increasing criticism, especially because the conclusions could often not be generalized: they were limited to the comparison of one specific or a few pictures with one specific text. In this study an attempt is made to arrive at generalizable conclusions concerning the relative efficiency of text, pictures and animations in user instructions.

To be able to draw general conclusions concerning the relative efficiency of text, pictures and animations in user instructions, the context of their production and use needs to be explained. Therefore, the use and need for instructions will be discussed first. Next, a short historical overview of the development of the user instructions will be presented. This shows some of the origins of present-day designs and the wide variety of instructive pictures, instructive texts, instructional elements (such as arrows) and the way these are combined. Also, the development of new media and its effect on instructing for product use will be discussed (chapter 2).

Chapters 3, 4 and 5 discuss the human-machine interaction on three levels, zooming in from an overall discussion of the designer-user communication process (3), via the level of user-help-product interaction and communication (4), to a discussion of perception and cognition in human-machine interaction involving user instructions (5).

In chapter 3 an analysis of the communication process is presented on the basis of the Shannon & Weaver sender-message-receiver-feedback model. This description indicates how complex this process usually is in practice. The sender is not a single user, but a team; the users may be millions with very different backgrounds; the channel may be much more than just the user instructions; the message may be very complex because the product is and the feedback to the designers is often minimal. This all should be taken into account when advice concerning the design of user manuals is based on the conclusions of research.

In chapter 4 the focus is on the communication between the user and the product. Here it will be argued that interaction and the communication between user and product may evolve along two lines: directly between user and user interface or indirectly when the user needs help to operate the user interface. Modern technology even allows for the product to be operated not via the ordinary user interface, but only via the help system (a test applying this is described in chapter 10). A model is introduced to account for the fact that in this communication the user has to switch up and down repetitively between the user interface and the help system. This has implications for the cognitive load and the mental model of users of complex products, as will be discussed in the next chapter.

Chapter 5 takes a closer look at some psychological aspects of the switching between user interface (interaction) and user assistance, in particular, the printed user manual or on-line instructions (communication). Taking goal-orientation and the hierarchical ordering of goals as a starting point of view, perception and cognition concerning the human-machine interaction are analyzed. It will be argued that perception and further cognitive processing of the user interface and of the user instructions may be very different. As a result, the development of a mental model may in such a situation be quite different from what is usually presented. A user who tries to operate a product purely on the basis of the user interface - excluding the user instructions - probably develops a mental model or device model of that product based on the graphical-spatial layout of the interface. User instructions, on the other hand, are much more linear-procedural in nature; therefore, a user who uses the user instructions intensively may well develop a more linear-procedural mental model, which will therefore be termed a 'procedure model'. It is up to the user to combine the possible different types of mental models to get an overall idea of how the machine functions and hypothesize how features should probably be activated.

Once the use, development and variety of instructions, and the practical context and the theoretical background have been elucidated, the development of the research concerning the use of various media in instructions will be discussed - including the increasing criticism (chapter 6). The taxonomy of information types in instructions, developed by Bieger & Glock, is then introduced. It is supposed that this taxonomy allows for generalizable conclusions. On the basis of this theory, a general hypothesis is formulated at the end of chapter 6. More specific hypotheses are derived from this overall hypothesis and tested in chapters 7, 8 and 9.

In chapter 7 the taxonomy of information types in instructions is applied for two-dimensional products. Bieger & Glock found that only four out of the nine types of information that they included in their taxonomy were really necessary in instructions (for three-dimensional products). In this chapter a test is carried out to identify whether indeed these four types of information are the ones that are vital for application in instructions for the use of two-dimensional products (displays, computer screens). Additionally, in this test some instructions presented in text are compared with those in representational pictures and in flowcharts (or 'logical pictures') on the basis of three types of information in user instructions.

In chapter 8 the relative efficiency of text and representational pictures and animations (based on the representational pictures) is compared. In this test only the operational information in user instructions for a telephone system is presented in text, in representational pictures or in animations (based on the representational pictures). Additionally, in this chapter, the use of presenting spatial information was tested: each of the versions in text, pictures or animation was presented with or without spatial animation (in pictures). All six versions were also compared with the existing instructions, which contained all types of instructional information.

Chapter 9 describes a test about learning efficiency as contrasted to immediate efficiency of instructions. This test repeats the test described in chapter 8, but now the subjects were asked to do the same tasks again, using exactly the same instructions in the same medium (either in text, in pictures or in animations and either with or without spatial information) one week later. This was done to study the memory effect of the various media used for instruction.

A different kind of test is described in chapter 10. This test deals with the possibility to automate on-line help systems, meaning that some types of information presented in an on-line help system may directly be activated by the user. This transforms the on-line help system partially into an alternative user interface for activating some features of a product. There is no comparison of the relative efficiency of text or pictures in instructions in this study. It is presented here because the application of various media for instruction should be considered with all possible contexts, one of which is that the help presented on a screen can in some situations be performed automatically by the system that presents the help (in this study a telephone with a display). In such a case the system presents not only the information, but also allows the user to let the system perform the actions needed to do the task.

In the Conclusions (chapter 11) the relative efficiency of text, pictures and animations is discussed on the basis of the findings in chapters 7, 8 and 9. This is followed by an attempt to answer the question whether indeed the applied taxonomy of information types may help to draw generalizable conclusions about the type of information that is necessary or desirable in the user instructions of complex products. Furthermore, an attempt is made to draw some conclusions for user-interface designers and user-assistance designers. Finally, implications of the findings are related to the more theoretical aspects, as discussed in chapters 3, 4 and 5. The findings will be discussed in relation to perception and cognition of instructional text, pictures and animation on the one hand and of user interfaces on the other. Also, the mental model is discussed as it is built up during interaction with products while consulting user instructions.

1.4. Definitions and descriptions

This study focuses on the presentation of information that users of modern technological products obtain from the user interface and from user assistance. These notions are complex and many different terms may be used for various elements of this information. Therefore, it may be useful to define or describe some of the terms in advance.

Preece (1994, p. 722) defines the 'user interface' as 'the totality of surface aspects of a computer system, such as its input and output devices, the information presented to or elicited from the user, feedback presented to the user, the system's behavior, its documentation and associated training programs, and the user's actions with respect to these aspects'. This definition is too limited to be used here, because it excludes non-computer systems. In this study 'user interface' will also refer to other technological products such as radios, route information systems, cameras and microwave ovens.

Preece's definition rightly includes documentation and training programs; too often these are forgotten. Dirken (1999, p. 270) has also stressed that product designers must feel responsible for the user instructions. In this study, though, for practical reasons, 'user interface' will *not* refer to the supporting information, such as user instructions, documentation and training programs (which is defined as 'user assistance'. This separation of user interface and user assistance is made here because in this study the focus is on the switching that users have to do between user interface and the supporting information in the user assistance (especially the user instructions, defined below) and on the use of various media in the user instructions.

For this study, 'user interface' only refers to the elements on displays and screens that can be used to directly influence the status of the product, or be a part of the process that influences the status: real or virtual buttons, icons, menus, indicator lights; and all graphic elements indicating grouping such as colors, lines, shapes, numbers and very short texts, dialogue boxes, feedback sounds, etc. All information that gives supporting information to help the user to install or use the product is called 'user assistance'. This may consist of a printed user manual, a reference card, extensive (professional) user documentation, a troubleshooting guide, an on-line help system (consisting of various parts), a guided tour, tutorial, a website with additional information that helps the user to use the product (e.g., with frequently asked questions), a helpdesk, training materials and sessions, etc.

It is sometimes difficult to decide whether certain information should be considered as part of the user interface or rather as part of the user assistance. Especially when interacting with software, users may receive windows that present both types of information, such as fill-in forms with descriptive information or tips[2]. Other examples are the small pop-up labels that present descriptive information about functions when a user moves the cursor over an icon ('tool tips') and 'assistants' or 'wizards': the parts of the software that guide a user interactively to

---

[2] For instance, in Microsoft Word one can insert symbols; if this is selected, a window appears with all kinds of letters and symbols. If a symbol is selected, sometimes the hotkey information is indicated (the hotkey is the combination of keys that immediately puts this symbol in the text). For instance, if one selects the symbol, on the Symbol-window appears the following information: 'Hotkey: Alt+Ctrl+E'.

a goal (like the installation of a printer driver). In this study all these types of information are considered as parts of the user assistance.

This study deals with user interfaces and user instructions for two-dimensional (2-D) products, like computer software, telephones, videocassette recorders (VCRs). Strictly speaking, these are not two-dimensional. Software, presented on a computer screen or display may have a completely 2-D presentation, but the keyboard and mouse are not 2-D. Microwave ovens, VCRs and other electronic devices are not at all 2-D. However, the control panels of these products can be considered as two-dimensional for users, as far as the instructions are concerned. The elements that have three-dimensional aspects are for instance the electronic interfaces at the back of the products where cables must be connected; these parts nowadays do not cause the major problems. It is the software installation and programming that cause the problems to be solved using the (almost) two-dimensional user interface and two-dimensional user instructions.

'User instructions', also called 'instructions for use', are the part of the user assistance that informs the users what the product and the various features can do for them and instructs them how to realize that. This information may both be on paper and on-line. It may contain a short reference card or a troubleshooting guide, but guided tours and tutorials are not parts of the user instructions. A 'short reference card' is a separate card or a few pages of the user manual that presents an overview of all features plus the combinations of buttons that have to be pressed to activate the functions. The short reference card contains no descriptions of the functions. The 'trouble-shooting guide' presents instructions that may help users who have encountered problems that cannot be solved. These instructions are often presented in alphabetical order, but they may be organized thematically, and very short troubleshooting lists sometimes seem to be presented without any order at all. The instructions usually start with an attempt to localize and specify the problem as precisely as possible and then inform the user what to do step by step.

The 'user manual' is the paper manual, a brochure or book that may contain the user instructions. The '(user) documentation' is the complete documentation with all information about a product, often technical documentation for professional products. Airplanes, submarines, mainframe computers and public telephone switches, for example, are sold including extensive technical documentation; it may consist of many big ring binders or even various CD-ROMs.

The 'on-line help system' or 'on-line help' basically consists of the user instructions that are presented on a screen or display. It may include more elements than user instructions on paper, for instance context-sensitive help. This information is dependent on the status of the device; usually this is only descriptive information ('What is this?').

A 'guided tour' gives an overview of the product by showing the major features; it does not give any instructions - apart from the instruction on how to activate the guided tour. The tutorial, on the other hand, not only informs about the product and its features, but may also contain exercises, so that the user gets acquainted with all features. A tutorial is a kind of personal training for the product.

CHAPTER 2

Product complexity and user assistance: an overview

This chapter presents an argumentation for the increasing need for user instructions as a result of the increase in complexity of products. Also, the development of user instructions for specific products on the basis of generic instructions that have been produced since mediaeval times will be shown to elucidate the wide variety of manifestations of user instructions that can be found nowadays. This overview will include a discussion of the introduction and application of various instructive elements, such as the arrow.

2.1. Products do not speak for themselves

Products seldom speak for themselves; we almost always have to learn how to use them. Only food products could be an exception. So, it may be in our genes to bite into an apple, but we have to learn that a chair is meant to sit on and that chopsticks are to put food in your mouth. For us, these products may seem intuitive, because we learned to use them when we were very young. But we need or have needed an instruction for every product that we use. We have learned to use many products at a very early age from our parents, without realizing that we are learning to use these products. Furthermore, the designs of products often give clues about the goals of the products and how to use them - they have 'affordances' as Norman (1988) explained. But even if these products have the 'right' affordances, they have to be learned: 'Every product is simple, but it has to be used in a certain way and each way has to be learned' (Norman, 1988, p. 26).

When we are confronted with a new, unknown product - a product with a new main function, or a new shape, it often appears to be difficult to even guess the function of this product. This is not just true for products from other cultures; it often appears to be difficult to guess the function of a new product from our own cultural environment, even if it concerns simple, non-electric ones[3]. Thus, it is even less likely that we would be able to guess how to use them. Someone has to show us, or we need printed instructions.

Until the Industrial Revolution, consumer products were relatively easy to learn to use. Learning to work with them was part of basic education. It was not until the beginning of the 20^{th} century that technologically somewhat more complex products started to invade people's personal environment. In fact, for most people, this invasion only started in the second half of that century. Moreover, since then, more and more buyers had to find out for themselves how to operate the products they bought. Resellers no longer had the time to explain the complete functionality of the products they sold to the buyers and an increasing percentage of the products were sold via mail order (Brockmann, 1998; Westendorp, 1998; Mijksenaar & Westendorp, 1999).

During the fifties, sixties and seventies of the 20^{th} century we not only saw the introduction of many new electronic consumer and office products, but also an increase in the number of features[4] that each of these products incorporates (Thölke, 1998). But they were all still relatively easy to use, as the number of features remained limited compared to present-day standards and there was always a separate button for each function. In the beginning of the eighties memory chips and microprocessors were introduced in a rapidly growing number of electronic products, not only in military and other professional products, but also in consumer and office products. As a result, many of these products have become very versatile and very difficult to use.

2.2. Towards more complexity

Especially in the last two decades, many new products have been introduced and each of these products may have dozens or even hundreds of features. Since the invention of the memory chip and the microprocessor, their performance has almost doubled every eighteen

[3] In an exhibition on visual instructions and in some classes, ten simple, non-electric kitchen appliances were demonstrated. For most visitors of the exhibition and for most students, it was impossible to guess the function of all ten products. Many people could not even guess the function of any of the ten products; nobody ever guessed all of the products right.

[4] 'A feature is an identifiable, product intrinsic, extra quality to a given standard defined by comparable products' (Thölke, 1998, p. 2). 'A feature is an extra quality to the product's core benefit' (Kotler, 1991).

months[5], and there is no reason to expect that this development will change over the next coming years (figures 2.1 and 2.2). At the same time, production costs and prices have been reduced and there is no reason to expect that this will change either. It is therefore reasonable to predict that the rapid growth of introductions of new products and new features will develop with at least the same speed as it has done until this moment. We will see the introduction of new versions of many computer programs, video cameras, telephones, microwave ovens, traveling and traffic information systems, stereo equipments, each provided with many more features than these products have now. Moreover, it seems reasonable to predict that we will see the introduction of many totally new products of which it is difficult to predict what they can do just on the basis of their appearance; this is a lesson that may be learned from predictions from the past (Corn & Horrigan, 1996).

It is sometimes stated that there will be a development towards products with fewer features (Business Week, April 29, 1991, cover article). Some companies have indeed produced and marketed simpler versions of products that they had on the market. A few examples make clear that these responses to complaints from the public and press were in vain. Around 1989, WordPerfect Corporation introduced LetterPerfect, a simple version of the famous WordPerfect 5.1 word processor, after complaints from users and in the press about the number of features in WordPerfect 4.2, 5.0 and 5.1. At about the same time, Microsoft Corporation introduced Microsoft Write for the Macintosh for the same reason. In 1990 Philips introduced the Easy Line, a series of audio and video products that had only the basic features. All of these products failed in the market and production stopped within a year. The conclusion: people complain about the complexity of the products they buy and they may scorn the variety of features that they 'do not use', but they go on buying them.

Thölke (1998) also recognizes this phenomenon. Following Kotler (1991), he argues that this is mainly because producers feel the need to come up with new products as often as possible. Because real new introductions are rare, Thölke states, producers introduce features - 'additions to the main functionality of existing products'. This

[5] Following Moore's Law, which gives an indication of the increasing performance of microchips.

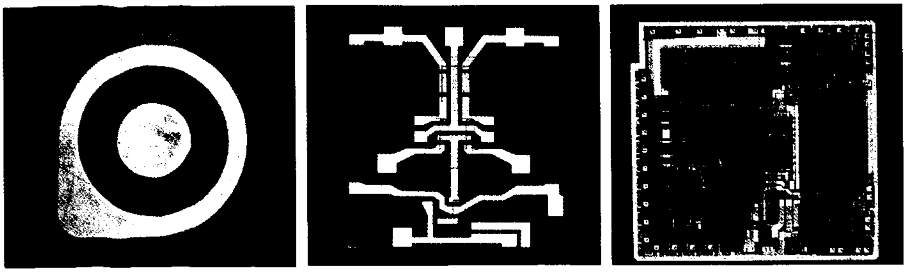


Figure 2.1.
Three examples of integrated circuits (Meindl, 1987, p. 102). Left: the first planar transistor (1959). Middle: the first consumer-oriented linear integrated circuit; it has five transistors (1964). Right: the first 16-bit chip, with an entire central processing unit; it has 20,000 transistors (1978). In 1959 the number of transistors on one single chip was 1, in 2000 chips with over 1 billion components have been realized.

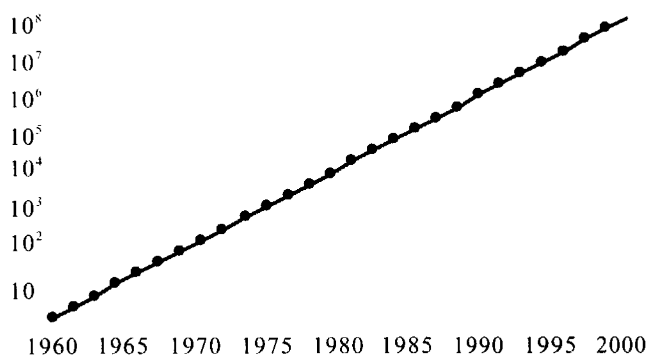


Figure 2.2.
Moore's Law gives an indication of the increasing performance of microchips. On average, every 18 months a new chip is released with roughly twice the computing power of its predecessor (www.intel.com/research/silicon/mooreslaw.htm).
Gordon Moore noticed in 1965 that up to that time the performance of the microchip had doubled every year. In later years the definition has changed (with Gordon Moore's approval) to reflect that doubling occurs every 18 months. Moore's Law is still remarkably accurate: in 26 years' time the number of transistors on a chip has increased from 2300 in 1971 (when Intel introduced the first microprocessor) to 7.5 billion on the 1997 Pentium II processor. The growth in performance of memory chips (important for the features in consumer electronic products) is just as impressive. 'Since the advent of the integrated circuit in 1959, the number of transistors that can be squeezed on a single chip has increased from one to several million. As a result, the performance of integrated circuits has improved by a factor of more then 10,000' (Meindl, 1987, p. 99).

conclusion is very likely to be true, but there are more aspects to explain the rapid growth of the number of features in products. Since the microchip has been mass-produced, prices have fallen sharply. As a result, it is relatively cheap to introduce new features in electronic products. This may by itself be a reason for consumers to buy the model with more features, rather than the model with fewer features. They may also reason that the features may not seem very useful right now, but may turn out to be handy later on. If the difference in price with the model with fewer features is not very big, they may easily choose to buy features that they may actually not consider to be useful at the moment they buy the product. Moreover, producers rarely market products in such a way that the consumer can choose for each of the features separately. Features usually come in groups: the model VCR that is just one step up in price may have five more features. So the consumer who wants just one of these features will buy the other features as well (Westendorp, 1993 and Westendorp, 1998).

Since the fifties, designers and ergonomists have been working intensively on user interfaces and on user instructions to make these products with all these features accessible for all users. The graphical user interface, the clustering and visual coding of buttons (buttons with comparable functions have the same color or shape), the textual and pictorial feedback via displays and via sounds, the design of user interfaces giving clues about the use of buttons ('affordances', see Norman, 1988), have all contributed enormously to make modern (electronic) products easier to use. But up to this day ergonomists, product designers and user interface designers have lost the 'competition' with electronics engineers and programmers: many more features were introduced than could be made easy to operate. As a result, these days, products are not easier to use than their predecessors of the past decades, but more difficult. This is easily seen in the development of typewriters (word processors plus printers), telephones, copiers, cameras and television sets. In the near future, the introduction of new features may even be much faster than ergonomists, interaction designers and user interface designers can cope with. "In the rush towards the digital era, we will continue to live right on the edge of intolerable frustration" (Odlyzko, 1999, Introduction).

An example may show this development and support the prediction. Programming VCRs has not become easier over the last twenty years, but more difficult, because of the introduction of new possibilities and features in parallel to the introduction of ergonomic aids. See, for instance, the entering of date, time to begin, time to stop, date and channel to record a specific program in the near future. When VCRs were first marketed to the general public, this programming had to be done manually, using the data in the radio and television guide. At the end of the eighties, Blaupunkt introduced a VCR with a bar code reader. Users could program their VCRs by scanning the bar codes for start time, stop time, date, channel, etc., printed on the last page of the user manual with a bar code reader connected to the VCR. It is questionable whether this system by itself made programming easier, but when radio and television guides started printing small bar codes next to the specific programs, this really seemed an improvement in the programming of VCRs. At the same time there is a development towards so many channels that a printed weekly radio and television guide with an overview of all available programs becomes almost impossible. Not much later, it became possible to program VCRs by means of a videotext system. This made the use of the extra bar code reader (the use of which had to be learned separately) superfluous and had several other advantages, such as topicality. But users had to be dexterous with the videotext system. The most recent development is programming via the Internet (see for instance Microsoft's electronic TV Host, www.tvhost.com). The latter makes it possible to use database queries so that one can let the VCR automatically record every program that has to do with a certain person or subject. But doing a database query via the Internet is probably not yet very easy for the majority of the target group[6].

It seems obvious that in a few years time for many people programming VCRs or recordable DVDs will be more difficult than today. At the moment this text is being written, digital television sets with a built-in Internet connection with an electronic programming guide are still a novelty. These television sets, of course, have features like authorized watching, preferred selection, auto-reminder, automatic recorder, programs not to miss and programmable filters, and many other new features in addition to the ones we already know. It is unlikely that these television sets and the accompanying VCRs will

[6] It is so 'easy' that only days after the introduction (1998) a small appliance was introduced that made programming via the Electronic TV Host really easy.

be easier to use or program than the ones that were sold thirty years ago. The time and effort users have to invest for the installation of these products has increased rather than diminished.

A recent development is that electronic consumer and office products look easy again, because they do not have many buttons. But the interaction problems for the user are not solved, just hidden: the difficulties are now hidden in menus, directly manipulatable icons and symbols, fill-in forms and many software buttons behind the screens. Remote controls and cell phones are good examples of such products that may seem easy to use because they only have a few buttons. It is unlikely that products are easier to use because they have a limited number of buttons. On the contrary, it is probably much easier to have a button for every function than to hide several functions behind one single button or under menus.

There are more arguments that prove that the difficulties people have with modern electronic office and consumer products have only just started. The user manuals that come with these products have become increasingly bigger with each new generation of these products. In four generations (in four years time), the user manual for a Canon laser printer grew from about 48 pages to several books, totaling some 750 pages. Moreover, new products come to the market with dozens or even hundreds of features right from the start. The Canon Wordtank is an electronic dictionary that includes an agenda, a scheduler and many other features. And the first edition has a 220-page user manual (per language). The calculator-sized gadget is nearly dwarfed by the instruction booklet!

2.3. Historical development of user instructions

The development towards a more general and more intensive use of technological products has its counterpart in the development of additional help.

Instructions in general may be as old as mankind can be considered a social and communicative species. Many gestures, facial expressions and cries may have been interpreted as warnings or threats and can thus be interpreted as 'instructions'. Human beings must have warned and signaled each other for danger or to help with hunting and other communal activities, both orally and visually. For thousands of years pictures and texts have been recorded on stone, wood, papyrus and other materials that may be interpreted as instructions. The Paleolithic paintings by the Cro-Magnon human in the caves of Alta Mira, Lascaux and other places in southern France and northern Spain (20,000-10,000 years BC) may be considered as purely descriptive, but can also be interpreted as prescriptive: instructions on how to hunt in groups (Boorstin, 1991, p. 894-895; Hering, 1997) (figure 2.3). Hieroglyphs in the Egyptian pyramids can be considered as purely descriptive (e.g., how a pharaoh was buried), but can also be interpreted as (prescriptive) instructions that indicate how a pharaoh should be buried. The Aztec paintings - as presented in the *Codex Mendoza*, an Aztec painted manuscript prepared on the authority of Don Antonio de Mendoza (1485-1552) (Ross, 1978) - may be just descriptions of the way the Aztecs lived, but the same kinds of pictures have probably been used, for centuries before the publication of this manuscript, to instruct the Aztec people about certain aspects of social life (figure 2.4). Likewise, in the *Kama Sutra* (Vatsyayana, 4th century AD) and in many other books and sculptures on Indian temples depicting the art of lovemaking, we can see many pictures that can be considered both descriptive and instructive (prescriptive). In Ephesis, Turkey, the sketch of a foot, a women's face and a heart in the pavement, dating back to ancient Greek times has been preserved, that is said to have served as a sign indicating the direction towards a brothel - descriptive (advertising) or prescriptive (road signage)? (figure 2.5).

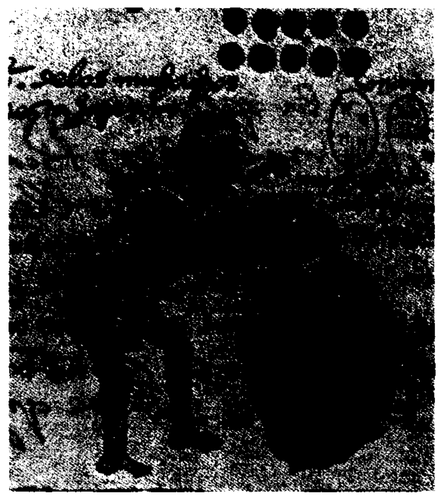


Figure 2.4. Illustration from the *Codex Mendoza*, showing how to punish a nine-year-old boy for a specific offense. The blue dots indicate the number of days that the punishment should be repeated (Ross, 1978, p. 78).

Moreover, there must have been technical drawings for the construction of technologically complex buildings that were erected thousands of years ago (Bremer, 1998). The Egyptian pyramids, for instance, cannot have been built without any instructive documentation. Probably, the first technical drawings were made for buildings. Since then, construction drawings have been produced for other complex technical products, like ships and windmills. Until the Industrial Revolution these products were all unique designs and for

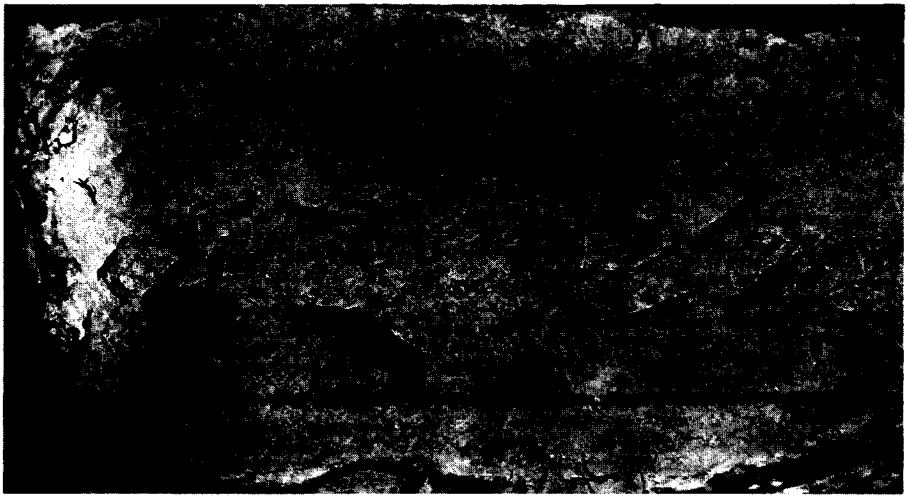


Figure 2.3. Paleolithical painting of a boar hunt on the wall of the Remigia cave at La Gasulla in Spain - about 20,000 B.C (Robin, 1993, p. 54).

Figure 2.5. **Footprint in the pavement at Ephesis, Turkey, pointing towards a brothel. Note the heart (or pubic hairs) and the face of a woman. Second century A.D. (Photograph: F. de Winter - van der Veer)**

every new ship new drawings were made[7]. The same was probably true for the construction of fortresses, windmills and other complex technological constructions.
In the European culture, from the Middle Ages onward, both visual and textual instructions have been produced (Gombrich, 1990; Brockmann, 1996). User manuals have been preserved from the 15th century onwards with some pictures that may be several centuries older. According to Gombrich (1990), many of them are strictly utilitarian, lacking any artistic ambition; others are very sophisticated instruction books. A typical example of the utilitarian user manuals is the *Fechtbuch*, a book on fighting, by Hans Thalhofer, dated 1443, which includes instructions on wrestling and unarmed combat (figure 2.6). A very sophisticated instruction book is *Wapen handelinghe van roers, musquetten ende spiesen* [How to use firelocks, muskets and spears] by Jacob de Gheyn; the original, secret edition dates from 1597; the public edition dates from 1607. The few hand colored examples of this famous book were for princes (figure 2.7). Both utilitarian and very well done is *De re metallica*, a book with 273 wood engravings on surface mining by Georgius Agricola (1494-1525); the complete edition of the twelve books was published in 1556; separate books have been printed since around 1530 (figure 2.8). Many fine illustrations can be found in medical books, the most famous of which is probably Andreas Vesalius' *De humani corporis fabrica* (or just *Fabrica*), dating back to 1543, with over 200 illustrations on 663 folio pages.

Two examples that illustrate movement are *Orchésographie* by Toinot Arbeau from 1589 and *Trattato della Piegature* by Giegher from 1639. The former is a handbook that uses the movements of fencing to explain the movements of dancing (figure 2.9). The latter is a treatise on the art of catering that contains visual instructions on the folding of napkins (figure 2.10), and may be one of the first instruction books to illustrate the exact positions of the hands in performing the task as well as the desired result (Gombrich, 1990).

Until the Industrial Revolution these instructions (just like the ones mentioned above) were generic, not specific. They instructed on horse riding, dancing, angling, fighting with fists, knives or swords and

[7] Curator of the Scheepvaartmuseum Amsterdam, personal communication.

Figure 2.6. (left) Hans Thalhofer (1443), *Fechtbuch* [Book on fighting], ed. Hergsell, Prague (Gombrich, 1990, p. 38). This book was recently published in English, see Rector (2000).

Figure 2.7. (right) Jacob de Gheyn (1597), *Wapen handelinghe van roers, musquetten ende spiesen* [How to use firelocks, muskets and spears]. This book was recently published in English, see Kist (1999). This reprint is from a color copy from the Delft Army Museum.

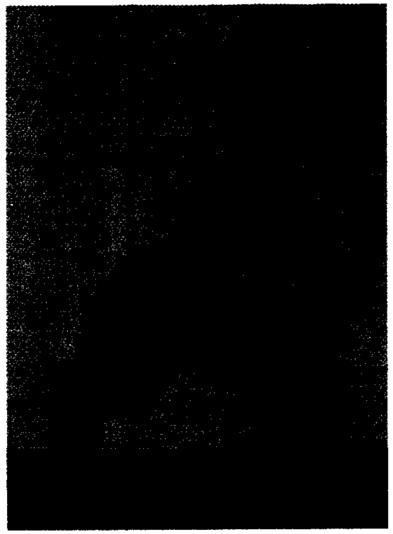


Figure 2.8. (left) Georgius Agricola (1556), *De re metallica*. An instruction book on surface mining. From the edition translated by Hoover & Hoover (1912/1950, p. 132).

Figure 2.10. (right) M. Giegher (1639), *Trattato della piegature*. Positions of hands shown (Gombrich, 1990, p. 41).

A—Upright forked posts. B—Pole over the posts. C—Shaft. D—First cord. E—Weight of first cord. F—Second cord. G—Same fixed ground. H—Head of first cord. I—Mouth of tunnel. K—Third cord. L—Weight of third cord. M—First side minor triangle. N—Second side minor triangle. O—Third side minor triangle. P—The minor triangle.

many other kinds of human activities, just like the books we have today on dancing, playing chess, swimming, brick-laying, making knots, cataract surgery and many other human activities. These instruction books and leaflets are generic, because they instruct how to do something with a certain category of products - or even without any product at all - and never for a specific product.

The only type of products for which specific instructive documents may have been produced before the Industrial Revolution are some military products, such as canons and other firearms. Brockmann (1998, p. 153) holds that the manufacturing of firearms was the only type of domestic technology that did not occur in the home or on the farm. Specific types of canons and muskets have been produced and sold to many armies in different countries. Some of these weapons

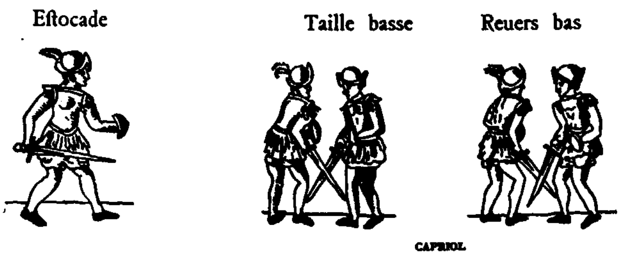


Figure 2.9. Toinot Arbeau (1589), *Orchésographie*. The movements of fencing are used to explain the movements of dancing (Gombrich, 1990, p. 40). The complete book is on the internet: http://memory.loc.gov/ammem/dihtml/dihowto.html

may have been accompanied by instruction books, but probably no example has been preserved. The oldest European book on the use of firearms is *Das Feuerwerkbuch* (figure 2.11), dating back as far as 1420, with handwritten versions preserved from around 1390[8]. This book is very much a generic user manual as it teaches the reader the art of making and using canons[9]. But in this book we also find examples of

[8] Nationalbibliothek Wien, Münchener Handschrift.

[9] 'Das Feuerwerkbuch von 1420 geht über das hinaus, was wir heute als Gebrauchsanweisung bezeichnen, da dort mehr als nur Anweisungen zum Gebrauchs eines Produkts gegeben werden. Das Buch beschäftigt sich mit der Kunst der Büchsenmeister im Sinne des lateinischen 'ars': das umfesst die Kenntnisse und Fertigkeiten in bezug auf die Herstellung von Feuerwaffen aller Art, Pulvermischungen, Geschoss- und Sprengmittelproduktion sowie den sachkundigen einsatz dieser Waffe im Kriege' (Plaumann, 1997, p. 55). [The Book of Fire-Arms (1420) comprises much more than what we would nowadays call a user manual, because it presents more than just the instructions for the use of a product. The book also tells about the art of the canon maker in the classical sense of the Latin word 'Ars': this includes knowledge and skills related to the production of firearms of all kinds, powder mixtures, the production of guns and explosives, but also how to use these at war.]

instructions for somewhat more specific products, like the use of a canon with the sight on the barrel or the use of a double canon with a semi-circular slanting disk (Münchener Handschrift 600).

Since the Industrial Revolution brought a great number of identical technological products to the general public, specific instructions became essential. Such user instructions instruct about the use of a specific product (many of the type may have been produced, but they are all identical). A generic instruction manual instructs on sewing, a specific instruction manual instructs on sewing with a specific model and type of sewing machine, or rather on just how to use this specific type of sewing machine (e.g., the 1853 Singer Mark I).

Figure 2.11. *Das Feuerwerkbuch* (c. 1420). In this case the generic instructions (instructing how to use a cannon) could also be considered specific (how to use this canon) (Plaumann, 1997, p. 54).

Brockmann (1998, pp. 153 and following) maintains that two types of products started the production of these specific types of user instructions: the sewing machine (since 1850) and the mower reaper (since 1850). Some comments can be made to this view. As Brockmann also concluded, of these two kinds of products, the sewing machine has probably been much more important for the development and production of user manuals for specific products. The mower reaper was sold only for professional use (farming) and to a much more limited public. Moreover, the mower reaper was only produced until the beginning of the 20th century. Brockmann states that next to the sewing machine, the automobile has been the most important product for the development of specific user instructions (he does not discern between general and specific manuals; his book only deals with the latter). A second comment is that at least the typewriters should be mentioned. The typewriter has rapidly become quite generally accepted and sold since its invention (Sholes and Glidden produced the first generally accepted typewriter in 1873; their machines were later produced and sold by Remington). In the last quarter of the 19th century many companies manufactured and marketed typewriters and sold them with specific manuals - although these user manuals usually also had a section on typewriting in general. In the early years of the 20th century not only the automobile should be listed as the product that introduced specific manuals to the general public, but radio equipment as well.

2.4. Presentation of user instructions

Instructions for specific products probably evolved from the generic instructive books, with influences from mapmaking (the use of symbols for real objects, the legend, the use of a scale, the arrows and warnings, e.g., for sailing routes) and scientific illustration (astronomy, anatomy, medicine and applied sciences, e.g., books about the calculation of the trajectory of cannon balls) (Boorstin, 1991, p. 149). Also, from the 18th century on, patents (often combinations of texts and pictures) may have had some influence. Indeed, the first specific user instructions (such as the 19th century user manuals for sewing machines and for typewriters) very much look like generic instruction manuals of the 15th century. Many early 20th century specific manuals also contain a lot of general information about the product; about what it can do, sometimes about the generic function (e.g., sewing, typewriting, making a telephone call or driving a car). The text is usually in prose and the pictures present the product as a whole, including all its details.

Lodges (1994) observes that many of the rhetorical features of the 15th and 16th centuries are similar to the practices of modern technical writing. For example, the *Treatise on fishing with an angle* (1496) 'makes a determined effort to relieve user stress about the new technology it introduces. It includes illustrations that supplement the text in ways similar to modern illustrations' (Lodges' formulation[10], p. 37). The treatise has seven illustrations and is referring to another one, which is not printed (so in this way too it is similar to modern user instructions!). Other modern features include the reference to illustrations in the text, the location of illustrations near the technical discussion and the use of call-outs to provide the names for the various parts of elements of the illustrations.

Hering (1997, p. 107) contends that there has been hardly any technological development from the Roman times until about 200 years ago (p. 107). He cites Booker (1964):

> 'Until 200 years ago, the materials of construction and energy sources available had remained substantially the same for thousands of years. The machines of war, corn-grinding mills, and water-lifting devices described in books of the late Middle Ages show only marginal advances on those that existed in the Roman times. Vitruvius' ten books on architecture, although written about 25 BC, were still used as text books by Italian architects in the 15th century'.

The idea that technology did not develop very much during the Middle Ages has been contradicted in recent years, and various important technological developments from the 16th, 17th and 18th centuries can be mentioned, such as the invention of telescopes, microscopes and other optical instruments, navigation instruments, medical instruments, printing machines, guns and accurate clocks (Boorstin, 1991[11]). We will not pursue this discussion here, because it is not the development of instruments, but the development of

---

[10] In articles written by writers or researchers of technical documentation, illustrations are usually considered as supportive to the text (Schriver, 1989 and the comments to this in Westendorp, 1994; Schriver, 1997); in articles or books produced by designers, texts are considered as a supplement to the illustrations (e.g., see Mijksenaar, 1997).

technical documentation that interests us now.

From the Roman times until the Middle Ages technical documentation may not have developed very much. Indeed, the present day generic instruction books, just like those of the 18th and 19th century, still very much look the same as those of the Middle Ages. But some developments should be noted. It may be worthwhile to have a closer look at some developments in the design of illustrations, texts, layout; and the use of instructive elements, concepts of user manuals and the use of various media.

2.4.1. Illustrations

The oldest preserved technical drawing dates back to 2130 BC. It is a representation of a drawing board from a statue of Gudea in Ur, now in the museum of the Louvre in Paris, showing the ground plan of a temple, including a measuring scale and a representation of a drawing instrument (Booker, 1964, p. 3). It is remarkable how little it differs from ground plans used by contemporary architects (Bremer, 1998).

It was not until the 15th century that the modern style of technical drawings, projection drawings with two-dimensional views, was introduced and widely used. Even Leonardo da Vinci, who has made so many variations of technical drawings, only made spatial (3-D) drawings (Booker, 1963, p. 127). It was not until the French Revolution that orthographic projection drawings became used. Gaspard Mongé (1746-1818) developed the basic principles of projection style technical drawing for the mass production of fortresses for Napoleon[12]. These principles were introduced in the United States in 1816 and were used regularly from then on. But it was especially the Industrial Revolution (when it became necessary to produce mutually interchangeable components) that made venerable technical drawings with very exact detailed information a prerequisite for mass production, and projection drawing appeared to be well suited (Bremer, 1998).

---

[11] This is a re-edition of Boorstin's famous book *The discoverers*, first published in 1983; the re-edition has hundreds of colorful examples of inventions and discoveries.

[12] Gaspard Mongé (1795), *Géométrie descriptive*. Paris.

Various artists have been credited to be the inventor of the perspective projection drawing. Booker (1963, p. 24) cites Giorgio Vasari, who claims that Paolo Uccello (1397-1475) was the inventor. But others are also mentioned[13]. Giotto clearly applied perspective projection around 1325; Alberti published his treatise on perspective *Della Pittura* in 1435 and Brunelleschi (1377-1446) invented a method of determining perspective projections in the early 1400s. But some ancient Greek vases show perspective illustrations. Albrecht Dürer made perspective projection drawing techniques generally known around 1520 (figure 2.12). These techniques were not adopted by engineers until the 18th century (Booker, 1964). That is remarkable, because painters and graphic artists adopted his ideas very rapidly throughout Europe and they also appear in instruction books.

Axonometric drawings originate in China, and were introduced in Europe in the 17th century by Jesuits. They noted how useful this drawing technique was for military strategy (artillery), cartography and technological purposes (gem cutting, diamond cutting, crystallography) (Krikke, 2000, p. 8). In contrast to perspective drawings, the lines in axonometric drawings do not converge at a (distant) vanishing point but remain parallel (Mijksenaar, 1991, pp. 7, 22-23).

Isometric projection drawings were introduced by Farish[14] in 1820 (Booker, 1964) (for examples of perspective and isometric drawings, see figures 2.13 and 2.14). 'Isometric' means 'of equal measure' and the isometric drawing method has made precise calculations possible, which is why this kind of drawing is widely used in industry. Isometric drawings are also very often applied in user instructions for laymen, probably because isometric drawings are easier to make than perspective drawings (Hering, 1997). Furthermore, this may be because the illustrations are derived from the technical drawings used to design the product, or because the same technical illustrator is asked to do the illustrations for the laymen. However, there may be good reasons to use isometric illustrations for laymen: 'Isometrical

---

13 www.cs.brown.edu/stc/summer/viewing_history/viewing_history_5.html

14 William Farish (1822), *On isometrical perspective*. Cambridge Philosophical Transactions. Vol. 1, pp. 1-20. Note: 'It is doubtful whether Farish actually 'invented' isometric projection - it probably existed in an empirical form among engravers' (Booker, 1963, p. 115).

Figure 2.12. Explication of the perspective drawing. Albrecht Dürer (1520), *The artist and the reclining woman* (Robin, 1992, p. 202).

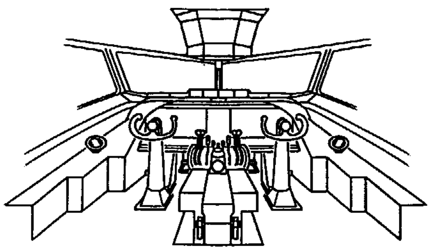


Figure 2.13. Perspective drawing. *Illustration guide for technical publications.* Fokker aircraft company (1987), Amsterdam.

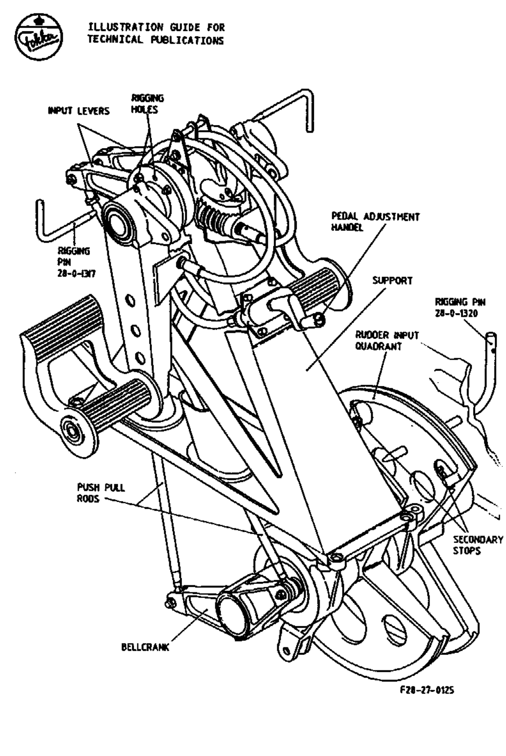


Figure 2.14. Axonometric drawing (isometric). *Illustration guide for technical publications.* Fokker aircraft company (1987), Amsterdam.

perspective, less faithful to appearance, is more faithful to fact' (the American architect Claude Bragdon, as cited in Krikke, 2000, p. 11). Bragdon even argued already in 1932 that the isometric perspective is 'more intellectual, more archetypal' and that 'it more truly renders the mental image - the thing seen by the minds' eye' (cited in Krikke, 2000, p. 11). Krikke argues that isometric drawings are widely used in computer games (like SimCity), especially those with a very high vantage point, because of the more dramatic effect - an argument that might be used to apply more isometric illustrations (with a high vantage point) in user instructions.
Cutaway drawings have been used in medical books at least since the end of the 15th century. Examples of cutaway drawings can be found in some books on medical instruction (e.g. Lorenz Phryes (Fries), *Spiegel der Arzny*, 1518 (figure 2.15) and Ioannis de Ketham, *Fasciculus medicinae* (1493) (figure 2.16). Agricola's *De re metallica* (1556) may be the first book to use cutaways for technical instruction. It has many fine examples of cutaway drawings showing men working underground and ghost view presentations showing parts of machines in mineshafts (figure 2.17).

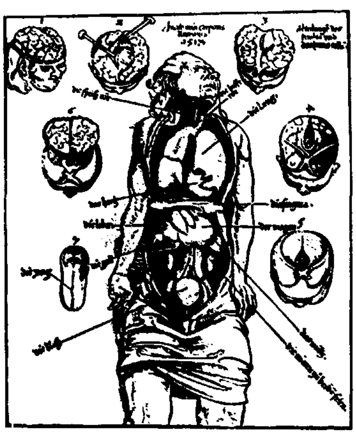


Figure 2.15. Combination of cutaway and exploded view. Lorenz Phryes (1518), *Spiegel der Artzny* (Herrlinger, 1967, p. 65).

Leonardo da Vinci may be the inventor of the exploded view drawings. He made many of these types of drawings to show the way products - either his inventions or ideas or existing products - were to be put together (Hering, 1997, pp. 107-108) (figure 2.18). However, there are quite a few examples from other drawings that very much look like exploded views dated even somewhat before Da Vinci's exploded views; see for instance the cutaway drawing by Lorenz Phryes (Fries) in *Spiegel der Arzny*, (Strassburg, 1518) mentioned above (figure 2.15).

2.4.2. Text

We may conclude that there have been some developments in the illustrations used in instruction books over the past five centuries. But the most important developments date back to the 15th century and the other developments have only been gradual. This is probably also true for the design and styles of texts in instruction books. Present day specific manuals seem to contain more structuring elements (detailed

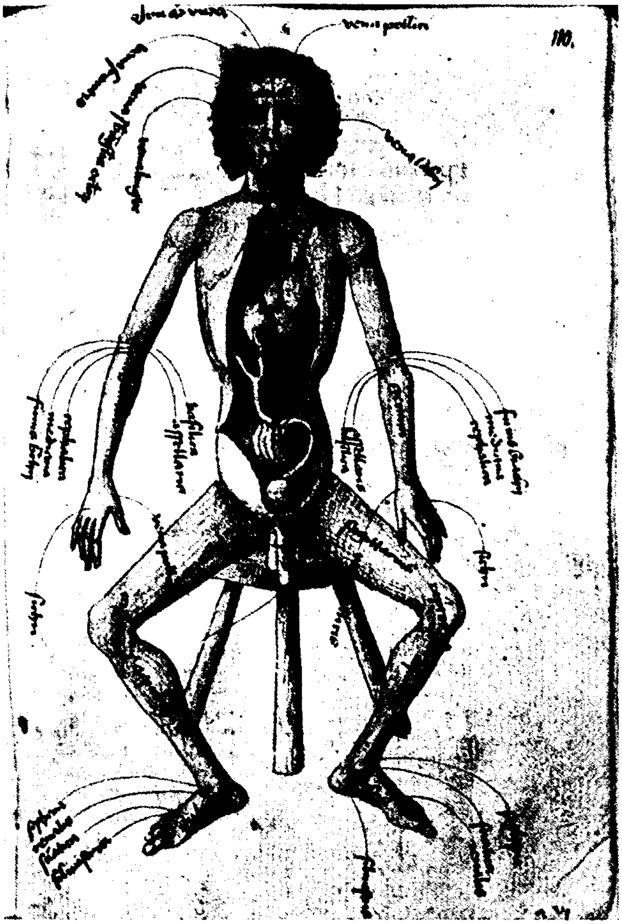
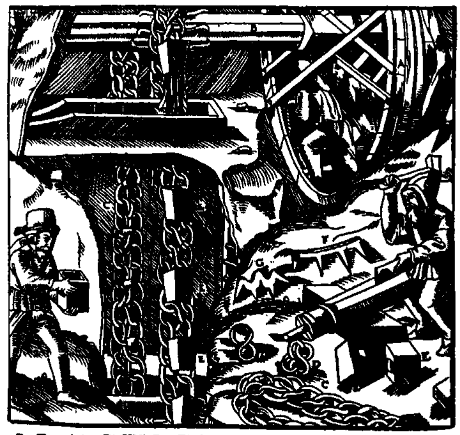


Figure 2.16. Cutaway drawing. Also one of the oldest examples with reference lines. Iohannis de Ketham (1491/1493), *Fasciculo Medicinae* (Herrlinger, 1967, p. 33).



Das Tretrad A. Die Welle B. Die doppelte Kette C. Ein Glied der doppelten Kette D. Die Kannen E. Eine einfache Klammer F. Eine dreimal gekrümmte Klammer G.

Figure 2.17. (above) Cutaway drawings. Georgius Agricola (1556), *De re metallica* (Hoover & Hoover, 1912/1950, p. 213).

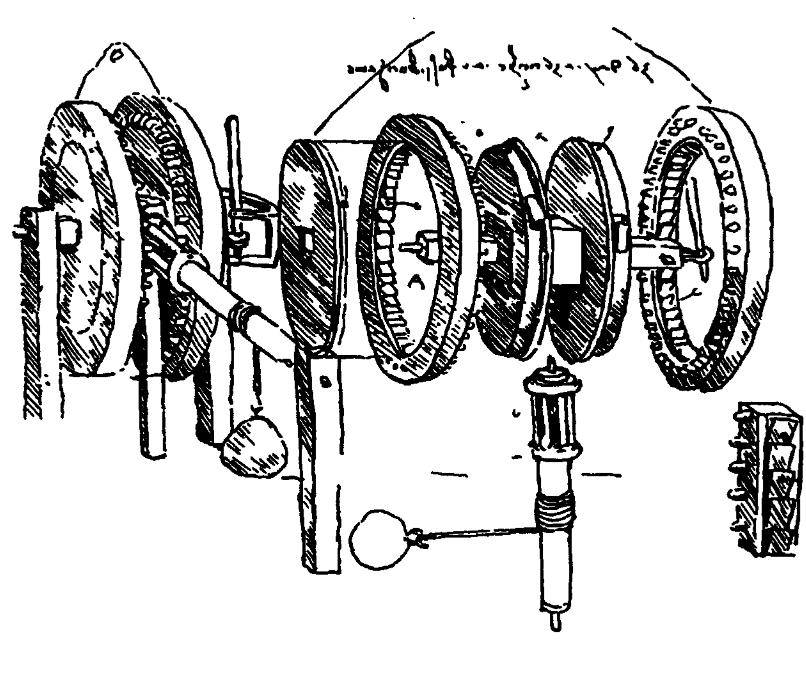


Figure 2.18. (left) Illustration of a winch by Leonardo da Vinci (1520). On the left the completed winch, on the right the exploded-view drawing (Hering, 1997, p. 107).

and numbered table of contents, headers and footers, index) and they also often contain lists of numbered short sentences. However, this is because we are comparing modern specific manuals with older generic manuals. If we compare the texts of generic, present-day instruction books with the generic instruction books of the past five centuries (e.g., cookbooks), the differences seem rather small - apart from the general differences in the use of the language.

Of course, the text style of technical user manuals of those days fitted within the culture of those days. Lentz (1998) discussed a technical user manual on shipbuilding[15] and concluded that it has all the characteristics of the classical rhetorical rules. For instance, the book opens with a dedication to the local government and sailing officials; the honorary gentlemen are praised very much and the author presents himself as a humble servant. The author addresses his colleagues in a separate preface and again begs the benevolence of his readers. Then follows a laudatory poem on the shipbuilding industry (by Sylvius). The actual book starts with an overview of the history of shipbuilding and the international shipbuilding traditions, but these texts are totally different in style: no longer narrative, but technical and purely instructive. Nevertheless, Van Yk, the author, also uses persuasive elements and anecdotes to stress the importance of the proper construction of specific parts. In an instructive, very concise dialogue style, the author is in direct interaction with an apprentice carpenter. This text style is not much different from the text styles of present-day generic instruction books; it may be just a little bit old fashioned.

2.4.3. Layout

Illustrations for instructions may not have evolved very much since the 16th century, texts may not have changed very much, but most striking is the wealth of layout variations in 15th and 16th century instruction books. Layouts may range from single overview illustrations and totally separated texts to central illustrations with integrated text or detailed

[15] Cornelus van Yk (1687), *De Nederlandsche Scheepsbouwkonst open gestelt*. Amsterdam. Gedrukt naar een origineel model uit de bibliotheek van het museum van Mr. S. van Gijn te Dordrecht. [The Dutch art of shipbuilding clarified. Printed according to an original model from the library of the museum of Mr. S. van Gijn in Dordrecht]. Published by Lagerveld, antiquarian at Rotterdam.

illustrations in the text or vice versa and many other variations. This variation in layouts of instruction books is probably not less than what we see in present-day generic instruction books (this may be slightly different for the specific manuals) (see figure 2.19, 1-6 and other drawings presented in this chapter).

Of course, there are instruction books without illustrations (e.g., many cookbooks) and there are books that mainly consist of (construction) drawings (e.g., for the construction of a certain building). But there are all possible kinds of combinations. Instruction books of the 15th, 16th and 17th centuries may have a text with separate pages for illustrations - either combined at the beginning or at the end or in the text, or a text with illustrations on pages in the text, or small illustrations integrated in the text pages. There are examples of instruction books with pictures with texts around them in captions or with either short texts or detailed illustrations in call-outs around the main picture. The texts may be in the pictures in all kinds of ways, for instance in boxes or not, upright or sideways, or along lines. Often illustrations have captions. Some books have separate pictures for the tools to be used and show the process of operation. Also, examples can be found of series of illustrations plus accompanying short texts showing the steps to be made (e.g., for surgery). Text may contain schemes and tables.

In the last century and a half, when user manuals for specific products were produced, a rapid development towards more illustrations in relation to text can be noted in the early years. After this sharp rise in the early years, the increase seems to have leveled off. Brockmann (1998, pp. 176, 179 and 238) has shown the rapid growth of the percentage of pages with illustrations in the (specific) user manuals for sewing machines. Between 1880 and 1915 the percentage of illustrations rose from 60 to 110. So, on average there was more than one picture on every two pages in 1880 already. In 1915 there was on average more than one illustration per page. A close analysis of Brockmann's data reveals that the number of pictures per page has sharply risen in a few years time, but after that does not rise any more. In fact, the percentage of pages with graphics in the Ford and Chevrolet manuals that Brockmann analyzed was a little bit lower in 1890 than in 1910. In 1860 the percentage of graphics in sewing machine and mower reaper manuals equals that for car manuals in 1910 (Brockmann, 1998, p. 238). In another study, Brockmann

Figure 2.19.1. Conceptual designs. Text plus pictures plus captions in the margin. Picture from George Silver (1599) , *Paradoxes of defense*, Campo, London (Coppens et al., 1998, p. 144).

Figure 2.19.2. Conceptual designs. Separate illustrations; referencing with characters. Picture from Nicolaas Petter, *Klare onderrichtinge der voortreffelijcke worstel-konst.* [Clear teachings in the art of wrestling] (Coppens et al. 1998, p. 171).

Figure 2.19.3. Conceptual designs. Captions in the drawing. Anonymous, 14th century Persian drawing (Robin, 1993, p. 169).

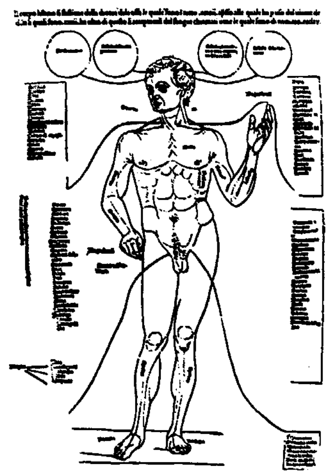
Figure 2.19.4. Conceptual designs. Overview illustration with referencing lines and captions in balloons. Ioannis de Ketham (Venice, 1493), *Fasciculo di Medicina* (Herrlinger, 1967, p. 38).

Figure 2.19.5. Conceptual designs, Overview drawing with referencing lines and detail drawings in boxes. Anonymous (1480). *Zodiac man* (Herrlinger, 1967, p. 31).

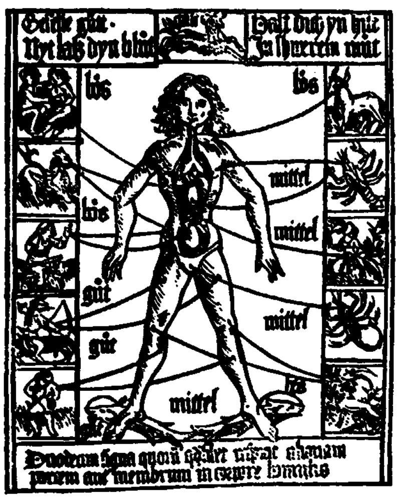
1

2

3



4



5

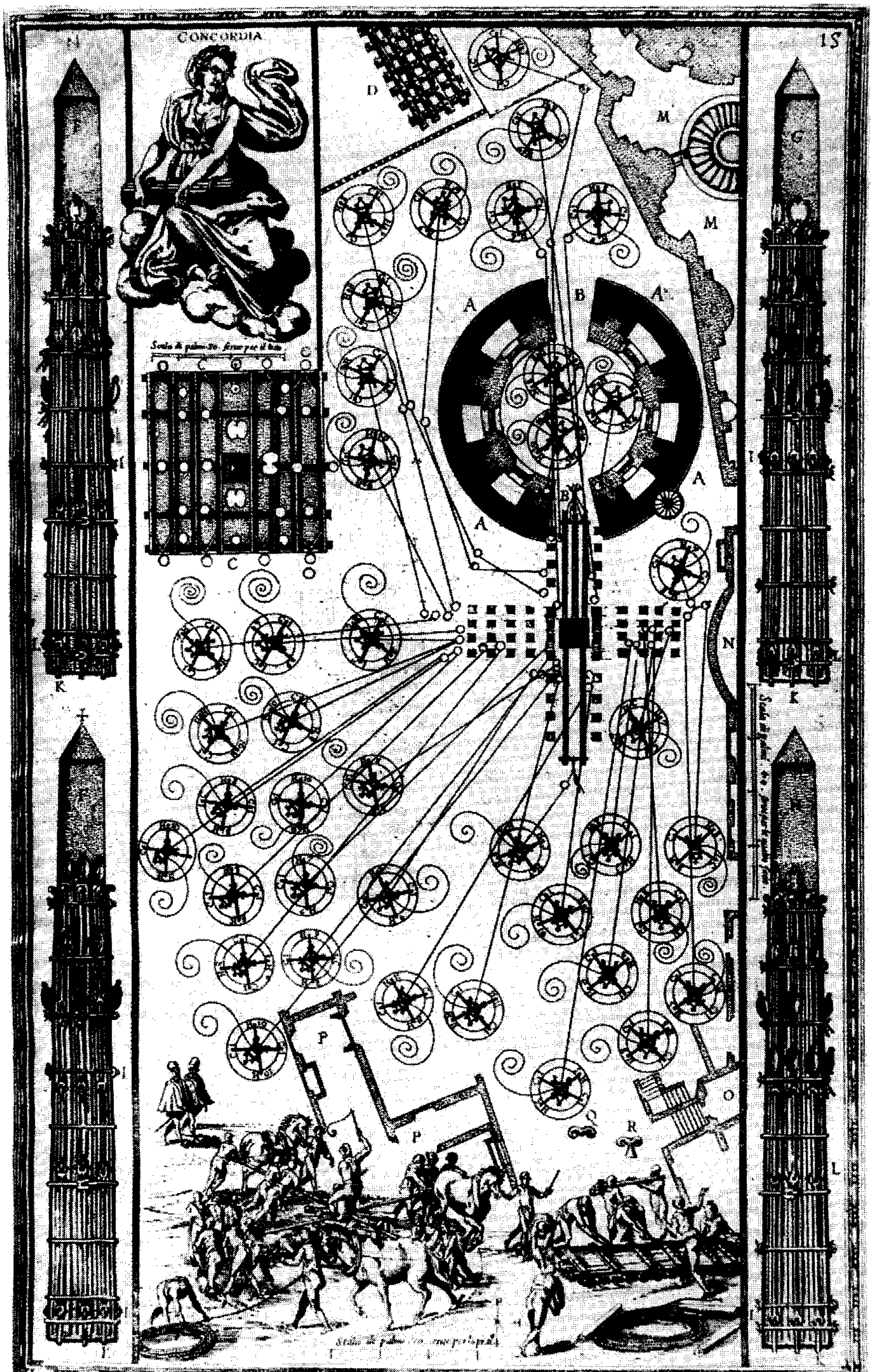


Figure 2.19.6. Conceptual designs. Overview illustration, separate from text. Overview with referencing lines towards enlarged details. Domenico Fontana (1586), *On moving the Vatican obelisk*. This instruction was not made for general purposes, but for the specific purpose of moving a 360-ton, 3000-year-old Egyptian obelisk from a square behind the Vatican and reposition it in front of Saint Peter's church at the Vatican. Fontana was in charge of this operation, which took 900 workmen and 74 horses between April and September 1586. Drawings of the plans were made by Nathale Bonifazio and Giovanni Guerra and printed by Bartholmeus Grassius to be sold as souvenirs during the operation. No instructions have been preserved from the transportation from Egypt and erection in Rome in A.D. 334 by Constantine (Robin, 1993, pp. 204-205).

analyzed Pagé's automobile manuals[16]. In Pagé's manuals, between 1913 and 1929 the number of illustrations in each manual rises, but the percentage of pages with illustrations drops from 70 to 48 (Brockmann, 1996, p. 291). A conclusion Brockmann does not draw is that the number and percentage of graphics in the manuals he analyzed rose sharply very quickly after the introduction of the new products (sewing machines and mower reapers around 1850; automobiles around 1910) and then remains about the same until the present day (with variations, but that may also have to do with the rather limited number of manuals that Brockmann has studied).

This extensive use of illustrations in sewing machine manuals (and in typewriter manuals) is continued in the 20th century automobile manuals. So, there is no development towards more illustrations in the period studied by Brockmann. The designers of the automobile manuals simply continued the style of the user manuals that was the state-of-the-art after fifty years in the sewing machine and mower reaper industry. 'The manuals of the automobile industry remained within the range of illustration profuseness established at the very end of the time period studied for the sewing machine and mower reaper manuals. (...) the 20th century automobile authors began writing within an established genre and understood that the manuals should have a large percentage of instructions and be relatively profusely illustrated' (Brockmann, 1998, p. 239).

The way pictures are used in the layout of manuals has changed greatly, especially after World War II. In the first century user manuals for specific products were produced (between around 1850 and 1950), overview-illustrations (especially at the beginning of the texts) were frequently applied. After 1950, and especially since around 1970-1980, a sharp rise can be noticed in the use of many small pictures representing the detail of the product to be activated or used. Pictures seem to be assigned another informative function: presenting just the operational information. This has led to a layout with columns of detail pictures next to columns with short text, a type of layout for manuals that is nowadays applied frequently, for instance for video and audio products and many other electronic consumer devices. Many

---

16 Victor Pagé was probably the most famous author of technical manuals in the beginning of the 20th century. His books were published by Henley (Brockmann, 1996).

manuals for cars and cameras seem to have multiple rows instead of multiple columns, but the principle is the same. In such manuals, pictures and texts seem to be supplementary to each other.

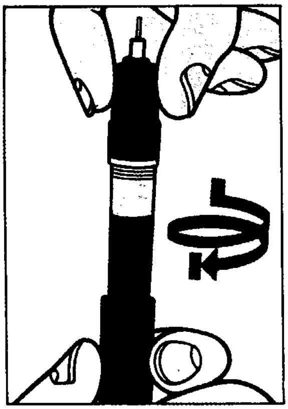
2.4.4. Instructive elements

An interesting development in instruction books and user manuals is the use of instructive elements in or near the drawings or as stand-alone 'language'. Instructive elements are additions to the illustration that do not depict any object, but are visual metaphors merely intended to instruct the user or help him or her to interpret the illustration. Instructive elements may range from arrows to crosses, traffic signs (introduced around 1900), hands, reference letters, numbers and lines[17] and other symbols (like the umbrella to indicate what side should be up or a skull to indicate danger, the hourglass to indicate 'wait'), textual onomatopoeia ('Click' or flash-like lines), colors (e.g., for localization) and anthropomorphic illustrations (where the product is illustrated with human-like characteristics, e.g., a sweating VCR) to indicate details (figures 2.20, 1-12). These elements in the illustrations are purely symbolic: there are no physical hands, reference letters, numbers and lines, arrows, crosses, dotted lines, exclamation marks, circles, zoom-lines or greyed-out or colored areas on the products[18]. It is interesting to notice that these types of elements, apart from the arrow, were introduced so much later: only after the beginning of the 20th century (traffic signs) and especially after World War II (when cartoonists were asked to help design instructions for the military). They indicate warnings, directions,

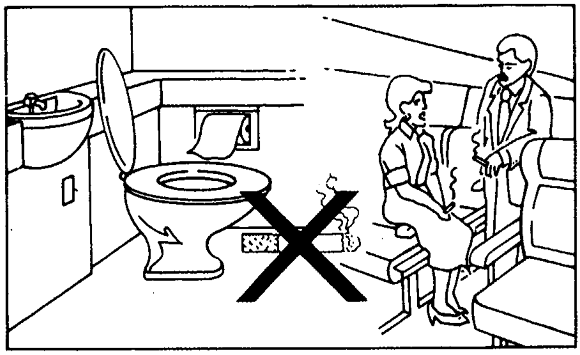
---

[17] 'Reference letters', 'reference numbers' and 'reference lines' are in a strict sense not instructive elements since by themselves they do not instruct. The arrow, for instance, may be an indication to 'turn left' or 'look at that detail', a picture of an umbrella may indicate 'hold upright', but a reference letter, a reference number or a reference line are not instructive by themselves. They may be interpreted as indicating 'look over there', so they are included in the instructive elements section. 'Hands' may be instructive and they often are, but they may be part of the scene that has been drawn (Mijksenaar, 1997, p. 44).

[18] Strothotte & Strothotte (1997, p. 54-56) discern realistic pictures and abstract-graphical pictures. The latter are 'pictures with properties of and relations in reality which are invisible to humans'. Strothotte & Strothotte (1997) also argue that 'verbs are awkward to express in presentational pictures'. In instructional pictures, the instructive elements ('abstract-graphical pictures') are in fact the verbs. These elements also make it possible to classify the picture as a whole as either true or false.

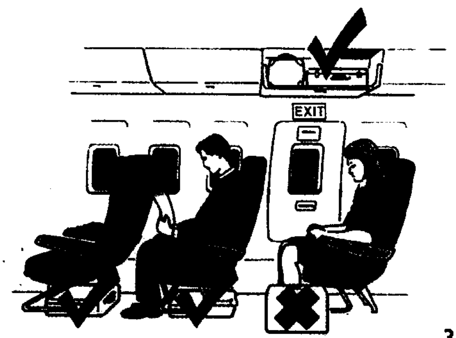


1

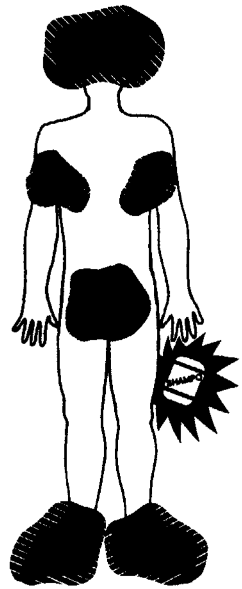


2

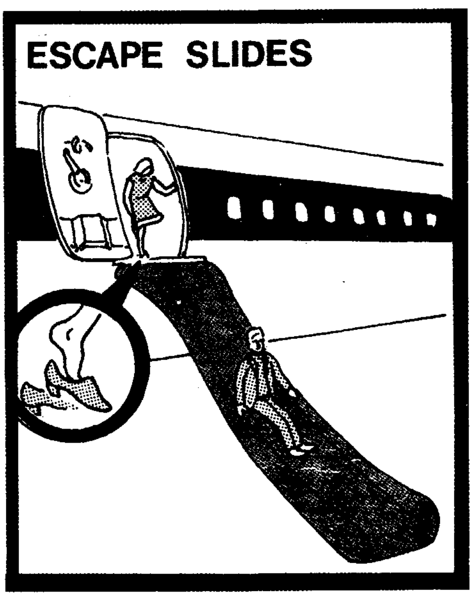
Figure 2.20. Instructive elements.
1. arrow to indicate movement;
2. cross to indicate 'do not';
3. cross and tick to indicate 'do not' and 'allowed';
4. red circle (prohibitory sign);
5. use of color to indicate area's;



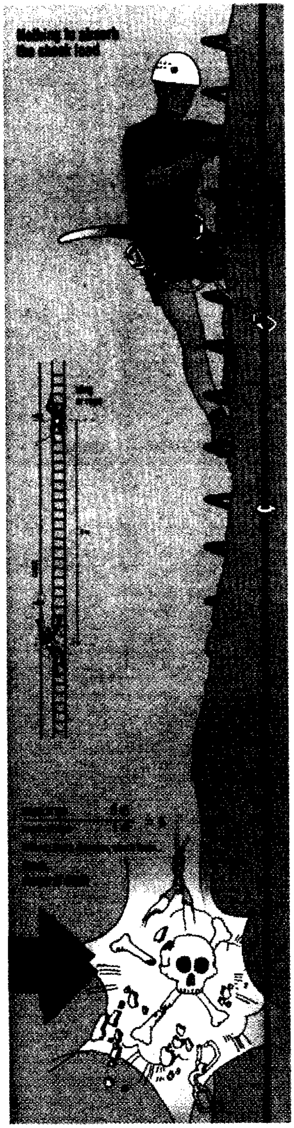
3



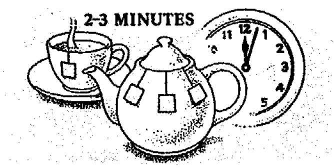
5



4

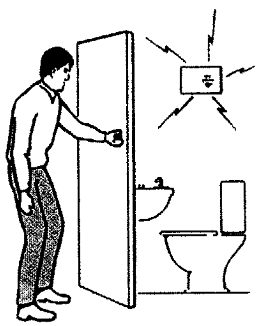


9



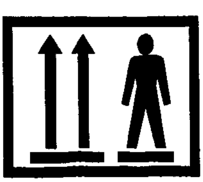
6

Figure 2.20.
6. use of color to indicate time;
7. flash-lines to indicate sound;
8. short lines to indicate light;
9. skull to indicate danger and arrow to indicate where to look.
10. arrows to indicate 'up';
11. glass to indicate 'fragile' (also used to indicate 'up');
12. umbrella to indicate 'keep dry'.

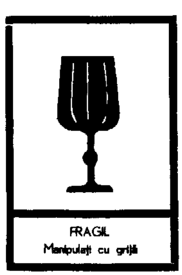


7

8



10



11

12

sequences, movements, connections or effects that actions may have. By far the most frequently used instructive element in user manuals is the arrow - which by itself may have various functionsincluding indicating direction, movement, location, connections and even sequences, but many other pictorial imperatives are introduced nowadays.

Instructive elements have rapidly evolved into a special instructive 'language'. The symbolic use of the arrow may be a couple of hundreds of years old (this will be discussed later) and traffic signs were introduced around 1900, but most other non-verbal, symbolic instructive elements in instruction books and user manuals originate in or after the World War II[19]. It is probably internationalization and the need of a generally understood limited instructional 'language' that evoked this development. The introduction of the computer for general use and the graphical user interface has greatly promoted the use of icons and symbols as instructive elements for general use (many symbols have been introduced for specific purposes, as in electronic engineering). Moreover, the use of computers and computerized products with screens or displays allows another development in the use of instructional elements: animation. Blinking of objects on the screen may be an example of this, for instance, when it means 'click here'. More developments of the language of instructional elements may be expected here.

In his interesting book *The symbolic species*, Terrence Deacon (1997) argues that humans are a symbolic species because and only because humans use (verbal) language (pp. 22, 25, 374, 375). It is questionable whether pictures by themselves cannot be considered as a symbolic language that other species do not have, but the instructive elements discussed here certainly form a language by themselves and therefore must be considered as another phenomenon that distinguishes us from other species[20].

[19] After World War I there have been attempts to develop a picture language that would be generally understood. This development finds its roots in post-War idealism that also brought Esperanto, Volapuk and other universal languages. Most influential has been Otto Neurath with his Isotype (International System Of Picture Education) language, a kind of construction box with hundreds of symbols especially developed for the graphical presentation of statistical data (Neurath, 1936; Van Mourik, F. & Mijksenaar, P., 1998, pp. 14 & 64).

(A) Hands

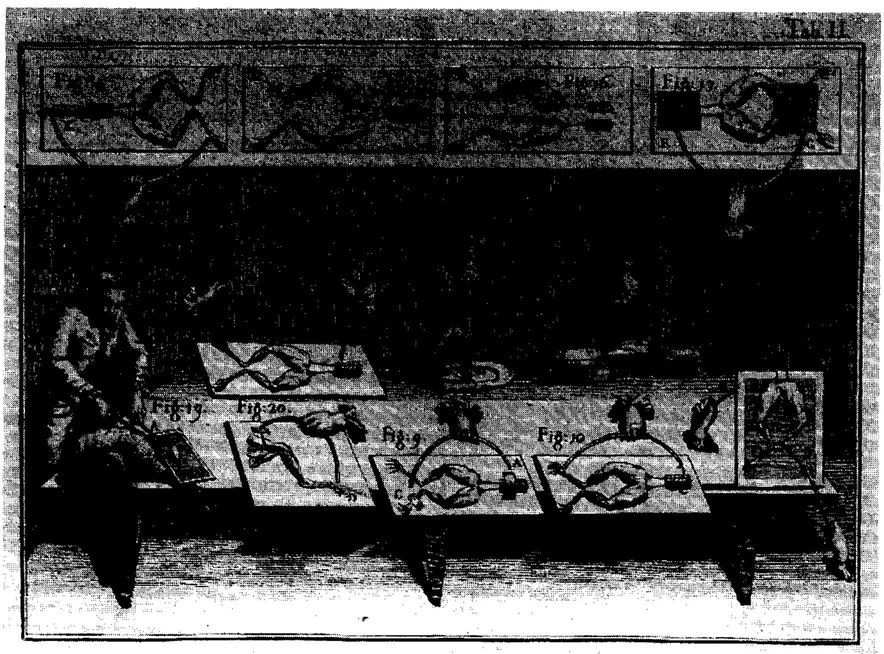
Hands in instructional drawings can be the illustration of the teachers' hands showing the apprentice where to look, so they could have been in the framing that the illustrator or photographer actually saw, but they are not a part of the product that has to be assembled. Therefore, they can either depict reality to show how the user had to hold their hands (figures 2.21, 2, 3 & 4) or also purely symbolic instructive elements (figure 2.21, 1). Gombrich (1990) writes that pictorial hands have been applied for indication since the Middle Ages and perhaps earlier. But it is very likely that the instructional use of a pointing hand is older:

> 'Although the pointing hand is more difficult to draw than an arrow, it is certainly a more primitive indicator of direction. The pointing hand is a gesture widely understood and it was early made into a pictogram. The precursors of the Chinese ideographs were called *ku-wans*, literally 'gesture pictures'. The Maya symbols for south and west incorporated hands outstretched toward the warmth of the sun and cupped in the position of eating. An Egyptian hieroglyphic hand represented the sound of 'd' and 't', but had no meaning as a pointer. [...] All this is to say that the pictorial pointing hand must be as old as man; but I cannot link it definitely with antiquity. The sign may appear in earlier manuscripts, but the first one recorded to have it is *Domesday Book* (1086) wherein, according to John Johnson's *Typographia*, it is used along with crosses, daggers, carots, and other marks, to indicate in the margin passages to be amplified or emended.'
> (McPharlin, 1942, p. 47)

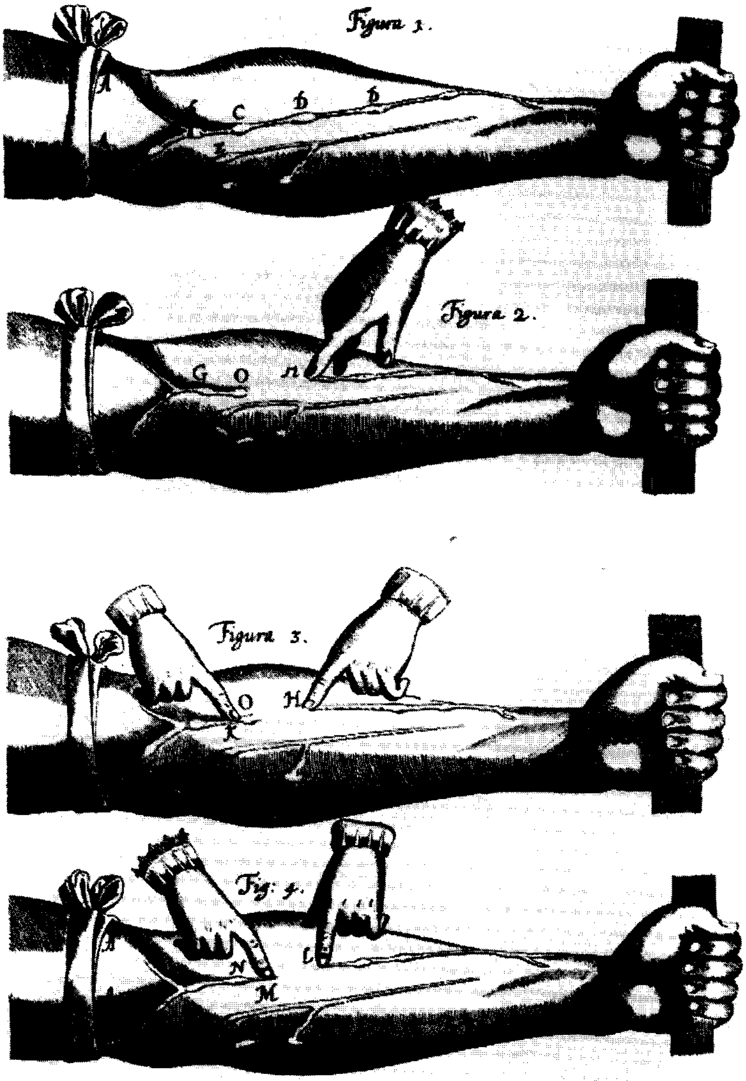
[20] It is unclear whether Deacon includes 'pictorial language' in his description of language when he argues that it is language that distinguishes us from other species. He does mention the possible non-vocal symbolic forms that language may have evolved from. 'The structure of syntax often only vaguely conceals its pragmatic roots in pointing gestures, manipulation and exchange of physical objects, spatial and temporal relationships, and so on' (p. 354). About the 'gestural origins theory', Deacon states: 'These facts suggest that gesture likely comprised a significant part of early symbolic communication, but that it existed side by side with vocal communication for most of the last 2 million years. Rather than substitutes for one another, gesture and speech have co-evolved complex relationships throughout their long and changing partnership' (pp. 355-356).

Figure 2.21.
Use of hands to indicate a detail (1) or to indicate realistically how to hold the hands (2,3). (1: Rawlinson manuscript, Bodleian Library, c. 1490) (2: Galvani, 1791). (3: Harvey, 1628). Example 3 also expresses a remarkable concept for visual instruction (Herrlinger, 1967, p. 19; Robin, 1993, p. 111 and p. 118).

1



2



3

McPharlin (1942) presents several 14th century examples of these kinds of pointing hands. They do not indicate direction, but focus special attention or indicate importance. Such pointing hands used in books have been very common for many centuries. They may originate in the Jewish tradition to use a silver pointer with a hand ('jad') to indicate lines of the Torah (which should not be touched with a real hand). The pointing hand or 'fist' sign often had a central position in the top row of the typesetter's upper case (McPharlin, 1942, p. 55).

For several ages hands probably have been functioning as indicators for direction, before the arrow took over this symbolic function, based on gesture (Aichler & Krampen, 1977).

(B) *Referencing letters and numbers*

According to Gombrich (1990), it was Da Vinci who introduced the alphabetical reference system (using this instead of labels with captions) to refer in the text to details in the drawings. Gombrich supposes that Da Vinci borrowed this reference system from Euclid, who used this in his books on mathematics. The alphabetical reference system quickly became generally used, many examples can be found in overview books of historic technical, medical, and scientific illustrations. The numerical reference system in illustrations was introduced around the same time. This system, just like the alphabetical system was widely used already before the beginning of the 16th century, making it less certain that Da Vinci was the (sole) pioneer with this system. For instance, Ludovicus de Prussia applied an alphabetical reference system in *Trilogium animae* in 1498 (figure 2.22). Ioannis de Ketham also applied an alphabetical system in 1512 in his *Fasciculo Medicinae* (figure 2.23), and Charles Estiennes and Estiennes de la Rivière applied a numerical referencing system in *De dissectione partium corporis humani* published in 1532 (figure 2.24).

It is remarkable that the alphabetical system has been applied much more often than the numerical system. Only very few drawings analyzed for this study used the numerical reference system, although in some cases this would certainly have been advantageous from a current perspective. It is not until the beginning of the 19th century that the numerical system is more generally adopted. But in the antique sewing machine and typewriter manuals the alphabetical system was used quite often.

(C) *Referencing lines*

Referencing lines have been introduced in instruction books at the beginning of the 15th century. Herrlinger discusses at length (1967, pp. 62-63) several hypotheses about the origin of referencing lines. He declines his first hypothesis that branding irons in drawings may have been their predecessors. Yet, his example with hands holding branding irons pointing at some parts of the body (figure 2.21, 1) seems convincing in showing that this type of illustration may have influenced some illustrators to use referencing lines. Herrlinger also

Figure 2.22. (left) Use of characters for referencing. Ludovicus de Prussia (1498), *Trilogium animae* (Herrlinger, 1967, p. 61).

Figure 2.23. (right) Use of characters for referencing. Ioannis de Ketham (1512), *Fasciculo Medicinae*, Antwerp edition (Herrlinger, 1967, p. 84).

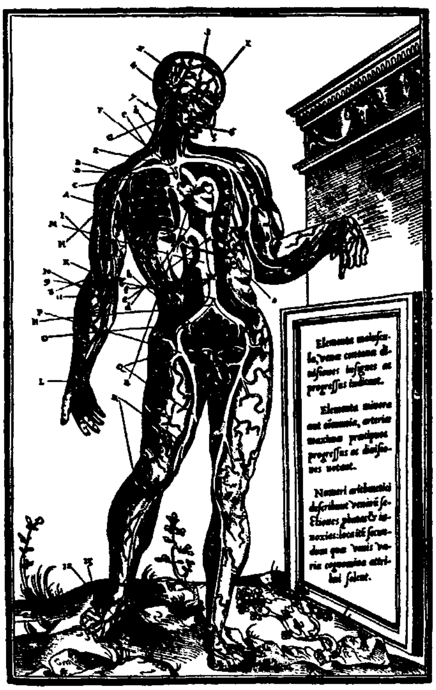


Abbildung 86 Tierkreiszeichenmann im Deutschen Kalender für das Jahr 1483 aus der Offizin KNOBLOCHTZER *in Straßburg.*

Figure 2.24. (left) Use of numbers for referencing. Charles Estiennes (1532), *De dissectione partium corporis humani* (Herrlinger, 1967, p. 95).

Figure 2.25. (right) Referencing lines. German calendar for 1483. *Zodiac man*, Knoblochter Office, Strassburg (Herrlinger, 1967, p. 62).

presents examples from 1483 with referencing lines between parts of the body and symbols of the zodiac (figure 2.25). He wonders whether the referencing lines may originate in the bleeding figures (showing where on the body to do the bleedings). However, he argues with some fine examples that the referencing lines probably originate from the use of 'gothic labels' (figure 2.26), which was at that time quite common in the sacral painting tradition. The oldest examples have loosely drawn lines, some with double lines that end in one point (so they look a bit like arrows); the wood engravings have dead straight lines. Nowadays, referencing lines are still widely used; sometimes they take the shape of an arrow.

(D) *Arrows*

The first symbolic uses of arrows are in compass cards, introduced by the ancient Greeks, perhaps Hipparchos. He introduced the idea to draw a notational imaginary network of meridians and parallels around the world and used arrowheads to indicate north on compass cards. Via compass card and compass (the first European mention is in the 12th century), the arrow was introduced in cartography (Aichler, & Krampen, 1977). In *De re metallica* (1556), Agricola presents some illustrations of instruments with indicators that have arrows and scales (figure 2.27, 1 & 2). Aichler & Krampen suppose that the first uses of arrows on maps - for instance to indicate the direction in which a river flows, date from the end of the 18th century. But this use of symbolic arrows may be somewhat older. Gombrich (1990) mentions Bernard Forest de Bélidor as the first to use an arrow in a technical drawing to indicate direction. In his *Architecture Hydraulic* (1737) Forest de Bélidor uses an arrow to indicate the direction of a waterwheel in one of his drawings (figure 2.28).

An interesting example of the use of an arrow is about 80 years older, but it is questionable whether it is a symbolic use. It is in a picture in *L'Homme*, a book by René Descartes reprinted in *The Scientific Image* (Robin, 1993, p. 71) (figure 2.29, 1). Herrlinger (1967) also refers to this picture, but states that it appeared (five times) in *Dioptrique*, a book published by Descartes in 1644 (figure 2.29, 2). The five pictures look very much the same, but there are minor differences so that they must have been produced separately. The picture copied here from Putschers' *Geschichte der medizinische Abbildungen II* [History of medical illustrations] (figure 2.29, 2) is indeed slightly different from the one copied from *The Scientific Image* (figure 2.29, 1). In these drawings arrows are shown,

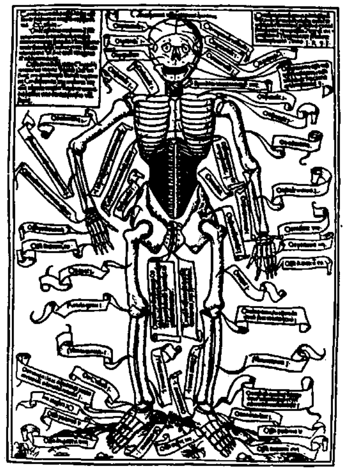


Figure 2.26. Referencing with labels pointing to details. Ricardus Helain (1493), *Ortus sanitatis* (Herrlinger, 1967, p. 63).

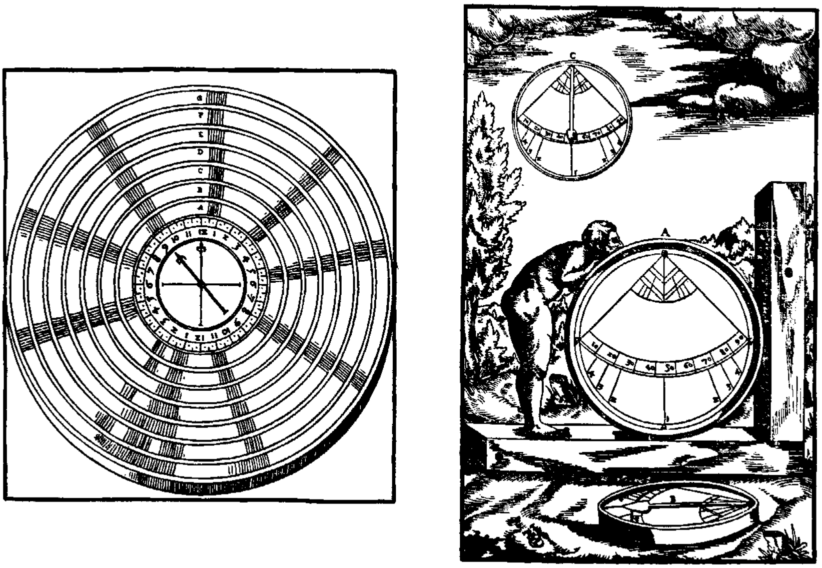


Figure 2.27. 1 & 2. Magnetic arrows on instruments. Georgius Agricola (1556), *De re metallica* (Hoover & Hoover, 1912/1950, p. 142).

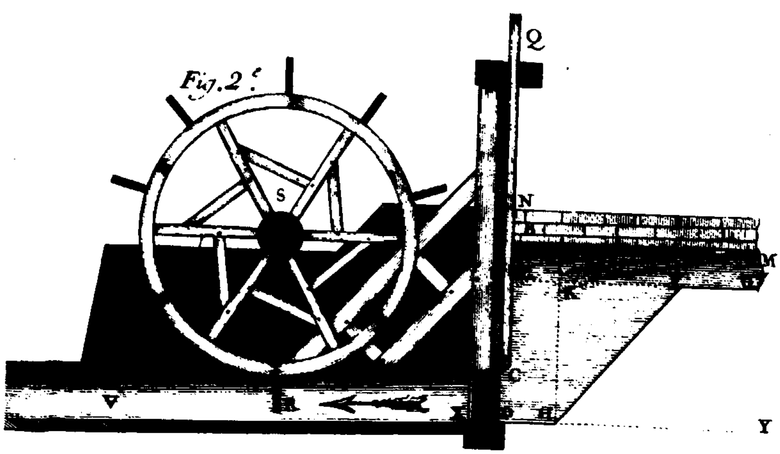


Figure 2.28. Bernard Forest de Bélidor (1737), *Architecture hydrolic.* Charles-Antoine Jombert, Paris. Diagram of a waterwheel. The first arrow found that indicates direction of an element Before that, symbolic arrows wre used on maps to indicate the rivers' directional flow (Gombrich, 1990, p. 28).

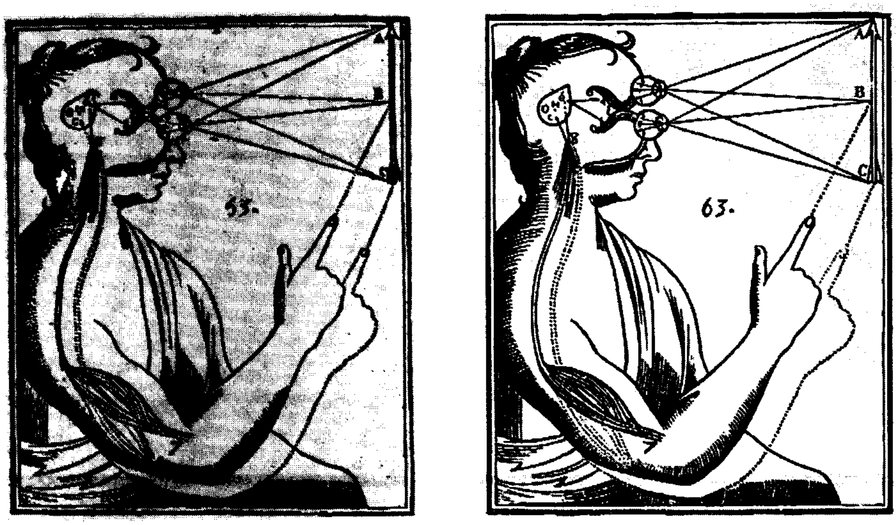


Figure 2.29.
Use of arrow. Use of dotted line to indicate alternative place of element.
René Descartes (1664), *L'Homme*
(1: Robin, 1993, p. 71)
(2: Putscher, 1972, p. 90).

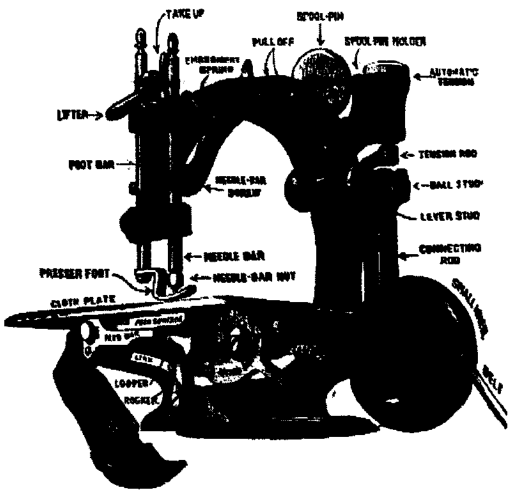


Figure 2.30.
Willcox & Gibbs (1869), Sewing machine manual.
Use of arrows.

but it is difficult to say whether this is only meant to be a realistic arrow that the person depicted is looking at, or whether it is (also) intended to show direction or distance.

Since the middle of the 18th century the use of the symbolic arrow has spread rapidly. In the second half of the 19th century arrows were sometimes painted on the wheels of sewing machines to indicate the direction in which the wheel had to turn. Arrows were certainly used to indicate elements of the sewing machine in the sewing machine manuals, as can be seen in figure 2.30, a detail from a 1869 Willcox & Gibbs sewing machine manual. Even before the 20th century arrows were introduced in traffic signs. The more realistic hand with a pointing finger, which had been used to indicate direction at least since the 17th century, was definitively replaced by the arrow.

(E) Other instructive elements

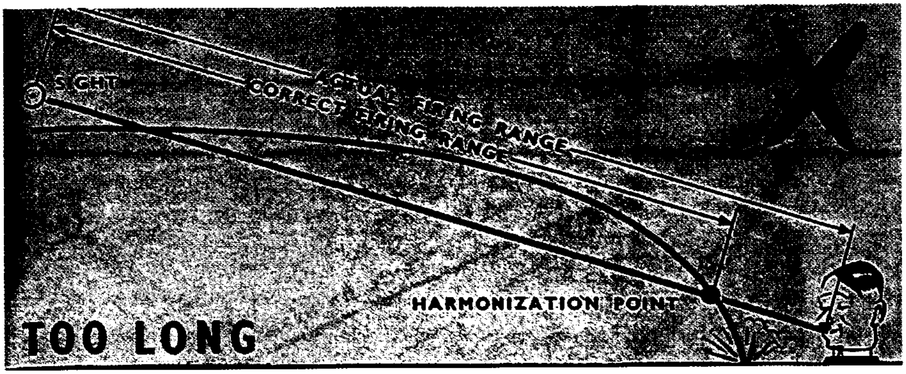
It is hard to find any instructive elements other than the hands and arrows in technical documentation - whether generic or specific - dating from before the 20th century. However, in this context it is worthwhile to have another close look at the illustration in the Descartes' book on visual perception (figures 2.29, 1 and 2.29, 2). In these pictures we see an arm drawn twice: once with a fixed line, just like the rest of the figure is drawn, and once more with a dotted line, to indicate a second position of the arm or the view from the second eye. This use of a dotted line is remarkable, since it may be the only example of this use of a dotted line to show movement or another position dating from before the 20th century.

Many of the cartoon-style instructive elements have been introduced in user manuals during World War II[21]. During this period many military specialists (like fighter pilots and gunmen) had to be trained in a very short period of time and the instructions had to be as simple as possible. It is since this period that cartoonists were hired for producing the illustrations for the instruction of pilots and other military men, and this relationship has had a lasting effect on the appearance of instruction books (figures 2.31, 1, 2 & 3 and 2.32, 1, 2 & 3 showing anthropomorphic illustrations). For instance, during World War II the Walt Disney Corporation produced a great number of instruction manuals, books and movies. They introduced their well-

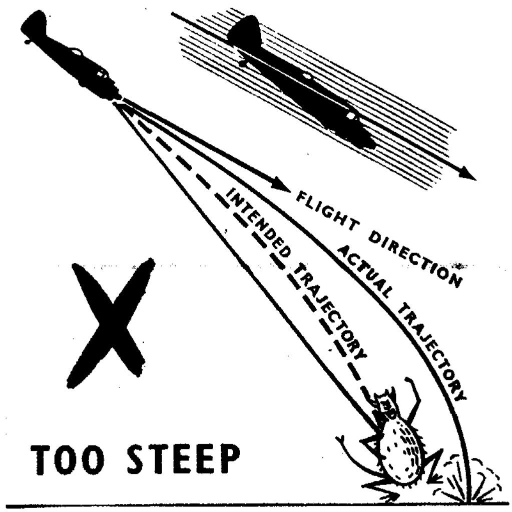
---

[21] Brockmann (1998, p. 156) disputes the notion that the instruction manual genre was born in the military manual mills of World War II. Brockmann wrote: 'Yet, like the guides and treatises for millwrights and shipwrights, the instruction manuals not only existed as early as the 1850s of Griffiths and McKay, but these instruction manuals self-consciously employed many of the design techniques now ubiquitous in the contemporary genre of instruction manuals including:
- dense interweaving of text and pictures
- packaging text into small paragraphs
- including indexes and tables of contents
- using task-oriented headings'.

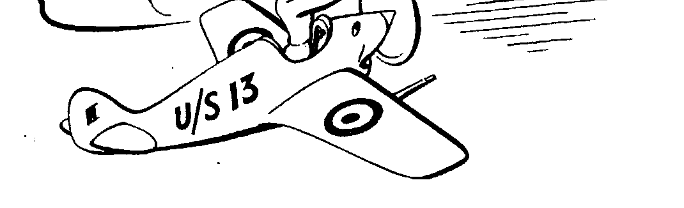
This may all be true (and Brockmann could have stretched this comparison as far as the instruction books from the Middle Ages and the Roman times), but anyone comparing manuals dating from before and after World War II will notice big differences, especially concerning the use of instructive elements in the pictures.



1



2



This method is definitely dated 3

Figure 2.31, 1, 2 & 3. Training manual for World War II fighter pilots. Cartoons. Note the use of green crosses to indicate what should not be done. *Bag the hun* (1943).

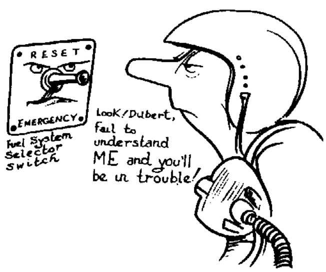
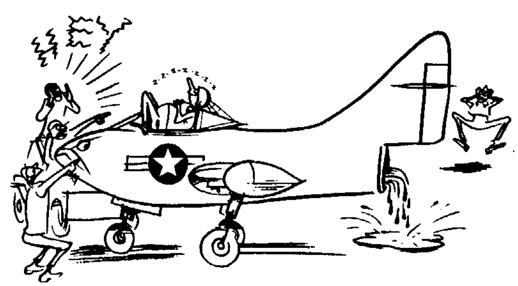
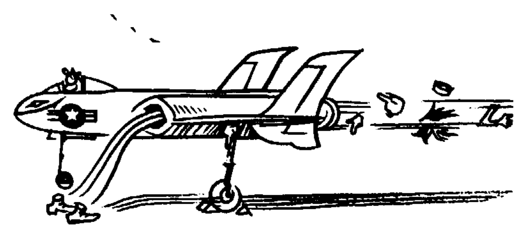


Figure 2.32, 1, 2 & 3. Training manual for early jet-fighter pilots. Cartoons. Note the anthropomorphic character of the airplane.



Dil doesn't chop the fuel the instant he recognizes he has a FALSE START.



known characters to explain the use of weapons. An example is the movie with Mickey Mouse explaining how to use the Browning P.50 machine gun[22].

Nowadays instructive elements are a major form of communicative help in instructions and they are often the instructive language by themselves. Look, for instance, at traffic symbols, handling instructions on packaging and warnings. Moreover, some elements have evolved to more abstract rules, rather than just statements comparable with individual words (Westendorp & Van der Waarde, 2001). This may indicate that a new symbolic language is evolving, which started many centuries ago, with the introduction of reference lines, call-outs, dotted lines and (later) the symbolic uses of arrows to indicate movement. The real development toward a new symbolic language that is or is supposed to be widely understood, however, is less than a century old, and the use of visual elements for instructions is rapidly evolving at present.

---

[22] Curator of the Dutch Army Museum, personal communication. The Walt Disney Corporation acknowledged the existence of this movie, but did not want to send a copy or photographs for publication.

2.4.5. Concepts

User manuals are introduced based on concepts that may vary altogether from the style of the 'standard' booklet with illustrations. The concept conveys more than just the layout on the page level and definitely also more than just the structure. Van der Waarde (1993) applied the phrase 'overall graphic presentation' to refer to all features of the graphic presentation that are related to a complete medical patient package insert (p. 50). This phrase is 'an overall descriptor of all the meaningful marks in a document and encompasses the spatial, graphic, and substrate features'. It is more general than other terminology used in graphic design research, such as configuration, layout, graphic organization or graphic structure, because these terms 'seem to refer more to the spatial arrangement of graphic components' (p. 50). Van der Waarde (1999) also presented a framework for the description of the graphic presentation of patient package inserts, consisting of three levels (graphic components, relation between these components and the global graphic presentation) and a total of eleven factors. For this study the phrase 'overall graphic presentation' and the framework are still too limited to describe the overall 'idea' that may be behind a user manual. Therefore, here the term 'concept' will be applied, although this term may be more difficult to define precisely.

Snoek & Hekkert (1998) define 'concept' as 'a materialized idea, in which the first signs of design are observable, but details are still lacking' (p. 172). This definition was set up especially for 3-D product design. In advertising and graphic design, 'concept' means the basic idea or setup of a campaign, of an advertisement, brochure or commercial. Floor & Van Raaij (1997) use and describe the word concept in this sense, but they do not give a strict definition; this appears to be the case in many books on advertising and design. In this book, the concept of a user manual is defined as 'the materialized idea in which the setup for the user manual as a whole is observable, but where details are still lacking'. The concept shows the way in which text and illustrations are related in a complete user manual; it is the basic combination, structure and presentation of the instructions in which the first signs of the design are observable but details are still lacking (Van der Waarde & Westendorp, 2002, and www.stic.nl, workshop April 20, 1999).

Some characteristics of user manual concepts are:

- basic solution to the instructive communication problem
- certain level of abstraction (overall design, main functions, no details)
- materialized (sketch or scheme)
- indicating how the user can install and use the product
- indicating the use of various media (text, graphics, animation, sound)
- indicating the combination of these media

The concept of a user manual is the answer to the question: in what way are we going to communicate this? In advertising this usually results in rough sketches of (series of) pictures with headings; the same may be the case for user manual design. The concept of the user manual should be enough to explain the basic translation of the communication problem. For examples illustrating what is meant here with 'concept' in the context of instructional design, see Mijksenaar & Westendorp (1999), chapter 6.

If we compare user manuals of VCRs in the same price category, with approximately the same features and complexity and user interfaces, we are comparing different utterances of approximately the same communication problem (Mirel, Feinberg, Allmendinger, 1991, p. 84). Yet, we may notice a wide variation of concepts that were behind the presentation, more than just differences in layout and structure. The designer may first of all decide to have a booklet or a broadsheet, foldout pages or combinations of these and, for instance, add a short reference card. Various languages may be combined on each page or presented separately. Of course, the formats may differ widely, from A4 or A5-size brochures in portrait format, to oblong (one having exactly the same size of the VCR and a full-size illustration of the VCR on each page). Some have a foldout page with an overview illustration and referencing numbers plus plain texts in columns (which can be used in combination with the various languages) or with texts plus detail illustrations. Others are based on detail illustrations and have separate texts or texts integrated in the illustrations. Again others have a separate sheet with all the detail illustrations plus a booklet with texts. To summarize: there is a wide variation of concepts behind user manuals, even for user manuals for the same type of products with approximately the same complexity.

2.4.6. New media

More recently, new media (sound and animation) can be applied. This started in the 1970s when VCRs found their way to consumers. The audio and - more importantly, - the videotape for instructional use have become quite generally accepted for generic instructions (e.g., flight training for pilots or laying bricks for amateurs), but not very much for specific instructions. There is one exception: many videotapes with specific instructions have been produced and marketed for learning to work with specific computer programs, such as AutoCad version 11, Lotus 1-2-3 version 3.1 or Microsoft Flight Simulator version 8.0. Philips produced an instruction videotape for their Video-2000 VCRs; this instruction was purely visual (no text at all), but a complete printed user manual was added. Instructions on videotapes were never a real big success. Videotapes do instruct quite realistically, but a major disadvantage is that they are basically linear: the information is just being shown for a certain amount of time at a fixed speed and order. The user can only sit back and watch, and at best stop and replay parts of the tape. Moreover, there is no interaction; the user is a passive spectator - not the best learning situation.

This changed rapidly with the introduction of instructions in the software (on-line help) on separate floppy discs, CD-ROMs or via the Internet. This software is non-linear (randomly accessible) and often has interactive instructions; the user is invited to perform certain tasks and the computer program checks whether the user has performed the right actions and may even give hints (e.g., in tutorials). The information and instruction are presented in combinations of media; animations and real video next to still pictures and texts and sound (either machine sounds or human voice).

The process of instructing users

The interaction between users and modern technological products is a complex communication process that involves many aspects. This study focuses on the design of the message, but to evaluate conclusions in the right context it may be important to discuss the communication process that designers of user instructions are confronted with. An analysis will show how complex this process is and, as a result, how difficult it is to generalize conclusions about isolated aspects (in casu the effects of media choice), and how difficult it is to draw conclusions that can be applied by user interface designers and user assistance designers.

Therefore, in this chapter we will discuss the communication process between the designer of instructions and the users on the basis of the 'sender - message - channel - receiver' model. This model seems useful to describe the complexity of each of the actors and elements involved in the process of communicating instructions for operating technical products.

After the description of this specific technical communication process, the focus will be on the human-machine interaction with information for the user via both the user interface and via additional user instructions (chapter 4). In on-line help systems where user interface and user instructions are presented on the same screen, this may lead to less difference between user interface and user instructions. The latter may sometimes even replace the former.

The difference in presentation of information, in the graphical-spatial user interface on the one hand and the linear-procedural user instructions on the other hand, may have serious implications. The user has to combine these two completely different types of information-presentations. Therefore, in chapter 5, some of the perceptive and cognitive aspects involved in these dual interaction processes will be discussed in more detail: perception, learning, memory, decision-making and, more specifically, the way we represent information and the cognitive load that various representations may have.

Chapters 3, 4 and 5 present a theoretical background to describe the situation within which the analysis and results should be seen. Just like chapter 2, these chapters are meant to indicate the framework within which overall conclusions must be judged. Moreover, an attempt will be made to draw conclusions from the discussion of this specific communication process to more theoretical psychological issues.

3.1. The communication process applied to user instructions

Conclusions that may be drawn from detailed studies concerning the relative efficiency of various media (chapters 7, 8 and 9) for informing users of complex technological products, transportation systems and administrative processes can only show their relevance once the entire process of information between designers of the information systems and users has been considered. To describe this information process, the well-known communication model involving **sender, message and receiver** (proposed by Shannon & Weaver in 1949, later introduced in the human communication research by Jakobson & Halle: see for instance Jakobson & Halle, 1956, p. 44), will be applied here. This is done just because it involves all elements that need to be discussed: the senders of the message, the content and design of the message, the way in which the message is transmitted, the characteristics of the receivers and the feedback from receivers to senders (figure 3.1).

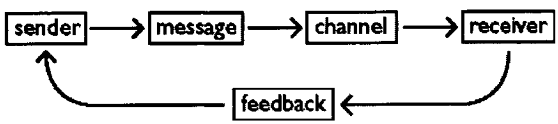


Figure 3.1. The basic communication model as introduced by Shannon & Weaver in 1947.

3.1.1. The sender

The sender of a technical instructive message is often not just one person. The messages that have to be transmitted are too complex, both in content and in design, to be made by one person. In this communication process the sender is usually a team that may consist of technicians, commercial executives, lawyers, graphic designers, writers, translators, illustrators, photographers and software programmers.

Technicians

The first people involved in the communication of the user instructions for technical products are the technicians, such as electronic, mechanical or software engineers. The information concerning how to use a technical product is often communicated via a multi-step-flow process: the engineers inform the writers or designers (and they instruct illustrators) and these inform the users. Of course, occasionally the engineers themselves may be the designers and writers of the user instructions. Both processes may have their advantages and disadvantages. Indirect communication may lead to loss of information or misunderstanding. If the technicians do not check the final version of the message, there may be mistakes because of this pre-communication process. On the other hand, if technicians design and write the user instructions for the end users who are not technical, it is very well possible that many things may be taken for granted, such as the use of certain terminology, the ability to read technical drawings, the functioning of some sub-systems, etc.

Commercial executives and lawyers

A second group of senders involved in the process of user instructions of technical products are the commercial executives: marketeers, marketing-communication managers and product managers. More and more often lawyers are also involved because of liability issues. For writers and designers of user manuals, the influence of these groups does not make it easier to communicate the main message: how to operate the product. The commercial department sometimes wants advertising and promotional information that the writers feel may interfere with the instructions they need to give. Lawyers sometimes want so many safety warnings that the really important warnings are obfuscated. Moreover, marketing managers or product managers tend to stress that the user manual has to be produced for less money. Most of them still do not consider the user instructions as an integrated part of the product they are selling, a part that needs careful attention and that could be an element in the marketing-communication process.

Graphic designers

In general, user instructions are designed by graphic designers. They usually devise the concepts of the user instructions, the layout and the graphic detailing. These concepts and layouts are often applied for long series of user instructions, sometimes even for a wide range of products. As a rule, writers fill in the texts and illustrators fill in the specific pictures within a fixed 'canvas'. The graphic designers may come up with ideas concerning the use of instructional elements and the choice of the type of pictures (photographs, various types of drawings); often they select the illustrators. Graphic designers may have a tendency to focus on the form, rather than on the content of the message; writers seem to do the opposite.

Of course, in instructive information, Horatio Greenough's 1853 statement 'form-follows-function' (made famous by Louis Sullivan in 1896; see: Mijksenaar, 1997, pp. 16-17) should be the first guideline. In instructing the use of complex products, the main goal should be that the message is communicated efficiently and effectively. Concept, layout, text and pictures must first of all try to convey what users should do. However, there may be some arguments for occasional predominance of 'form' over 'function'. User instructions have to be opened and used first of all. Now, the main reason to do so will of course be that the user is confronted with a problem that cannot be solved by trial and error. But often producers hope that the buyer of a new product will start reading the user manual[23]. In order to get the users to read and use the manual, it should not only be efficient and effective, but also attractive.

Within this context one specific ongoing design discussion should be mentioned, because it has such important design influences. The basis of this discussion is the 'Less is More' phrase by Ludwig Mies van der Rohe. With this statement, Mies van der Rohe meant to say that all elements that do not contribute to the design (message) should be left out[24]. In user instructions, the 'less is more' statement could be interpreted as a plea for efficiency. 'Less is more' can be considered the

---

[23] It is remarkable that they do not often ask the user to first read the on-line help system.

[24] This is comparable to Occam's razor, often cited in discussions about methodology in science. Occam's razor is a principle attributed to the 14th century logician and Franciscan friar William of Occam. The principle states: 'Entities should not be multiplied unnecessarily'. The most useful statement of the principle for scientists is, when you have two competing theories that make exactly the same predictions, the one that is simpler is the better. Occam's razor is often cited in stronger forms than Occam intended, as in the following statement: 'If you have two theories which both explain the observed facts, then you should use the simplest until more evidence comes along'. This principle goes back at least as far as Aristotle, who wrote 'Nature operates in the shortest way possible'.

fundamental statement of the Bauhaus movement (Heskett, 1987, pp. 91-113). This statement may be important for both writers and designers of user instructions; it was also basic to John Carroll's minimal manual ideas (Carroll, 1989). Designers who take this principle as a starting point for the design of (for instance) a user manual will try to leave out all elements that are not really necessary to convey the message. This may concern content and presentation, both texts and pictures, and the way they are combined and structured.

Some architects and designers have reacted to this strict Bauhaus doctrine. It was the architect Robert Venturi who reacted with the statement: 'Less is a bore'. He argued that the first task was to make buildings and products attractive to use. The Memphis school of designers was a result of this movement (Bürdek & Drukker, 1996, p. 244 and following). For user instructions, the 'less is a bore' statement could occasionally imply that not efficiency, but attractiveness is the most important aspect. Users should enjoy using the instructions and this is not necessarily just based on efficiency[25]. In their designs of user instructions, designers make decisions involving the 'efficiency first' or 'attractivity first'. Of course, in practice, there is no opposition, but rather a scale of possibilities that the designer will choose from. It is, however, important to realize that, on such a scale, designers always must take a position between 'efficiency' and 'attractivity'.

It seems that many researchers do not realize that attractiveness or style may be considered a functional aspect in instructing the user. In many contrastive tests, researchers have used pictures that were first of all completely arbitrary, and secondly, often considered to be bad or ugly pictures by art directors and designers (Van Mourik & Westendorp, 2000). Several art directors and designers have commented in this way about the pictures used by Schriver (1997), Peeck (1987), Strothotte & Strothotte (1997) and others. Bieger & Glock (1984/5, 1986) also did not account for the arbitrariness of the style of the pictures they used in their studies on the relative efficiency of applying either text or pictures for various types of information in user instructions. It is

---

[25] Attractiveness in Memphis style could be the use of funny ornaments, such as the notorious paperclip figure used in Microsoft Word help - or other just-for-fun illustrations in the help system. Attractiveness is probably more functional, if applied following the Form Follows Function model, when focusing on typeface, color combinations, layout and other design aspects that influence readability and attractiveness.

also hard to find discussions in research literature about the influence of style and aesthetics in user instructions. Dwyer (1972) does mention and illustrate the possible effects of differences in style; Peeck (1989) does comment about it, and Nordbotte & Crosby (1999) describe this factor for a related field of interest (the effects of style on the interpretation of statistical data).

Illustrators and photographers

In the process of instructing users, illustrators and photographers may play an important role. They may be specialized in instructional technical drawings and photographs, but often they are not. Their main interest may be making cartoons or drawings or photographs for advertising, where aesthetics may have a different function than in instruction. In practice, the illustrators are often the people that come up with ingenuous details, e.g., to indicate movements, location and orientation of an element of the product, the sequence in which actions have to be performed and the reactions of the product.

Writers

Writers of technical instructions vary widely in backgrounds, ranging from instructional science or linguistics, former teachers or copywriters or journalists. Many others have a technical background (especially if the writers are employees of the companies that produce the products) or have had some other kind of position in the company, ranging from secretaries to programmers, from marketeers to administrative personnel. And, indeed, sometimes the writer is simply the one who has not much to do at the moment the user manual has to be produced, just like Robert Pirsig described in 1973 in his novel *Zen and the Art of Motorcycle Maintenance*. As a result of this wide variety of text writers for technical documentation, texts may be very different in content, design and style - also when made for the instruction of various, nearly identical products. Some writers focus on the technical correctness, some on the linguistic correctness, while others may stress the educational effectivity.

Translators

Of course, writers have most influence on the text, but translators are not machines either. They may translate a technical instructive text with a certain personal touch. First of all, a translation can never be completely one-to-one, not even in technical instructive texts. Even in these kinds of texts, the connotation varies in different languages. But personal ideas about the text influence the translation as well. A translator into Italian of an international company insisted that in Italian it was impolite to use the imperative and the *tu* form (users should be addressed in the *Lei* form). Now this is clearly a personal opinion, maybe influenced by the fact that this translator was rather old (over sixty-five). Other translators and native Italian writers of technical documentation were not conscious of this. According to these much younger translators and writers the imperative and the *tu* form were both certainly acceptable and indeed desirable in user instructions. The sub-cultural background (possibly caused by age of the writer and translator, and because languages, including technical instructive ones, do change over time) caused the difference of opinion concerning the style of the text, not the cultural background (the fact that the translator is Italian).

Software programmers

Software programmers may have a strong influence on both the design of the user instructions and even on the text. On-line help systems are often designed by the programmers alone, and in these cases they are the ones who decide upon the structure of the on-line help, the way it is integrated or not integrated in the product, the design of the screens, navigation, etc. They may set limitations to the text as well, for instance concerning the length.

3.1.2. The receivers

It will be difficult to find a type of document that has a wider variety of readers than user instructions for a widely sold consumer product. The target group of user instructions for a radio, television, digital camera or word processor may vary widely concerning all thinkable aspects of a target group, such as age, education, experience with technical products and with the type of product, technical skill and acumen, intelligence, attitude, motivation, character and nervousness. Probably there are many other factors that influence the way in which the target group will perceive and use the information provided in the user instructions (easily being distracted, reaction to environmental situation, etc.).

For consumer products, the sender usually has to communicate with a target group that is much more varied than other target groups in mass communication. Journalists and copywriters, for instance, may have a target group of millions of people, but their audience is much more specified by education, income, interest and other factors characterizing the target group. Moreover, journalists know that their readers are to a certain degree interested in their articles - and copywriters know that their audience is not a priori interested. Designers of user instructions do not know this: some readers may be interested, some may not be. Readers of journalistic articles are interested in the articles (otherwise they would not have bought the newspaper) and readers of advertisements are not interested but attracted, but the readers of user manuals are neither interested, nor attracted: they are forced to pick up the user manual, because they have a problem with the product they have bought. They are interested in the product, not in the user manual or on-line help system. As far as motivation of the readers is concerned, the producer of user instructions is more aptly compared with a writer of schoolbooks than with journalists or copywriters. Schoolbooks, however, have narrowly described audiences (e.g., age, level of education, culture, language).

For many modern electronic products, the target group of user instructions is so varied that it is almost impossible to make the instructions understandable for all users. For the senders of the message it must be a frustrating idea to know in advance that many users do not want to read what they wrote and will not be able to understand what they want to explain. It may be impossible to explain even the basic functions of the products to all users. In the Netherlands it is estimated that about 7% of the population is functionally illiterate. This means that they can read, but only with much difficulty. They cannot read rapidly enough to follow subtitling on television. For these people, reading the user manual of a VCR, microwave oven or mobile phone will be too difficult. It is also often stated that many older people hardly even try to read the user manuals (Van Hees, 1996).

Moreover, the users of products may feel a certain tension when consulting the help. They have to get a new product working and they may be afraid that they will not succeed or even destroy or damage the product. Also, the installation or programming of a new product is often felt to be a kind of intelligence test. Journalists, copywriters and authors of schoolbooks are not faced with such receivers of their messages. In conclusion, designers of instructions for the use of modern electronic products have the most complex target group imaginable. For some products this means almost everyone.

3.1.3. The channel

In the 'sender - message - receiver' model the difference between 'message' and 'channel' is not always easy to distinguish. In the present situation, for instance, it is difficult to decide whether the 'navigation' is part of the 'message' or part of the 'channel'. One could also wonder whether in this model 'context sensitivity' is part of the 'channel' or part of the (presentation of the) 'message'. As this model is only used here as a general guideline to describe various aspects of the communication process concerning the instruction of users of modern technological products, issues concerning the model will not be discussed here. Some elements could just as well be discussed in another category of the model; the important thing is that the most relevant aspects are mentioned, to make sure that the complexity of these aspects is described.

We will begin by looking at the channel, the way in which the message is transferred. It must immediately be stressed that all user assistance (user manuals, on-line help, tutorials, guided tours, websites), but especially the on-line help should be considered as a part of the user interface[26]. If technological products would not have a dedicated user interface, one would have to be an electronic or mechanical engineer to be able to operate them. The user interface is designed to make products usable for ordinary users - so the user interface can be considered as the first part of the help system to operate products (figure 3.2). But modern technological products - especially the ones with silicon - may have so many features that it is impossible to operate them without additional help. Therefore, these products have two different channels to reach the user: the user interface and the user assistance (user manuals, on-line help, tutorials, etc., but especially the user instructions) (figure 3.3).

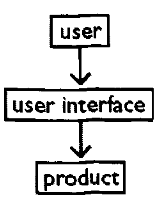


Figure 3.2. User-product interaction via the user interface only.

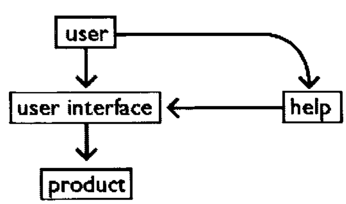


Figure 3.3. User-product interaction via user interface and supportive help.

[26] Donald Norman (1988) and Lewis & Rieman (1994) consider user documentation as part of the user interface. Dirken (1999, p. 270) has noted that product designers should not just design the product and user interface, but also the user instructions, so that they may themselves discover shortcomings in the user interface, or parts that are too difficult to explain.

The user instructions may just be supportive (as they are in most cases) to the user interface, but in some software it is already partly an alternative to the standard user interface. In these programs the user can occasionally operate the program via the on-line help system. A user who asks the help system about the installation of a new printer driver may get the possibility to install the printer within the help system; in some Microsoft software products, for instance, wizards that guide the installation can be started within the help system (figure 3.4). In the near future, the interactivity of the help systems (meaning that one can execute functions directly from within the help system) will probably increase (Hoek & Kaufman, 2000). Of course, usually, the interaction between user and product passes both via the user interface and via the help system as shown in figure 3.5.

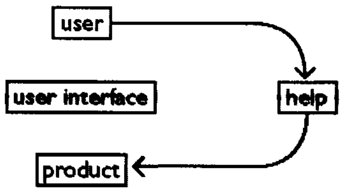


Figure 3.4. User-product interaction via the help system only.

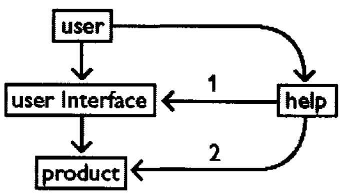


Figure 3.5. User-product interaction via user interface and help system. The help system may support the user interface (1) or be an alternative interface (2).

The help message is transferred via a combination of various channels. Therefore, the various channels and types of instruction materials that they convey must be discerned. The user's first information about the way the product functions may come from the vendor: advertisements and brochures. It is already at this moment that the user develops a mental representation of the product, even without even once having seen it. Occasionally, the brochures already contain some instructional information (e.g., if the seller wants to show how easy the operation of the product is). Other information about the possibilities of the product may have come from friends, relatives or neighbors. There may be instructions on the packaging and on the product itself (hardware). The most common types of user instruction are the paper or electronic documentation (technical documentation), the user manual or on-line help system (usually based on contents, index and find). But it is quite common to include a tutorial (lessons, including tasks), a guided tour (a demonstration of the major functions of the product), tips, and a reference list (a short overview of how all functions can be operated, without explanation). Moreover, intelligent products may produce information on the basis of user actions (feedback, such as error messages, but also pro-active information, based on user's actions) that can include instructions. Additionally, help information may be presented in 'frequently asked questions' on websites and discussion groups and via help desks. Some users may have taken special training courses for certain products. Usually, the help system for modern products, especially for software products, consists of a combination of all these modes of help.

Users can develop mental representations of the products they buy and use through various ways (figure 3.6) (Westendorp, Jansen & Punselie, 2000, pp. 19-21):

1. On the basis of what they already know or suppose about the product (a priori model).
2. Directly via the user interface (user interface model).
3. Indirectly via all user assistance channels: user manual, on-line help, guided tour, tutorial (user assistance model).
4. Training courses and other personal after sales information (training model).

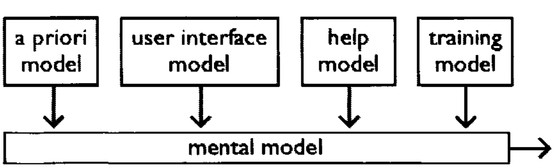


Figure 3.6. The user has to develop one mental representation of the product on the basis of various input mental models: via 'a priori information', via 'user interface', via 'help information' and via 'training information'. 'Help information' and 'training information' are nowadays also classified together to form the 'user assistance'. The arrow below right represents the continuous development of the mental representation during use of the product. For a more detailed version of this diagram, see Westendorp, Jansen & Punselie, 2000, p. 20 and back cover.

As the information in the user interface and in the user instructions are presented in different ways, the mental representation of the functioning of the product may be built up in different ways. One could even speculate that a user may build up various different mental representations of the same product, just because of the different ways the instructional information is presented. This will be discussed in chapter 5.

The second distinction that we can make is between the use of paper or an on-line presentation. The relative advantages and disadvantages of one medium over the other have been discussed extensively and with conclusions varying over time (e.g., because of the development of computers and screens). For an overview of these discussions, see Bouwhuis (2000); for a discussion of the integration of paper documentation and on-line help, see Duffy, Palmer & Mehlenbacher (1992) and Horton (1994). These discussions will not be repeated or analyzed here; it may be enough to observe that on-line presentation is gaining territory over paper presentation rapidly. On-line documentation is cheaper and can be produced and updated faster, but especially improvements of computer technology, computer screens, user interface design and of the design of the user interface of the help system have made this change possible and will make further developments in this direction possible.

3.1.4. The message

The message in instructional help may be very complex, both in content and in presentation. For many modern technological products, the content is complex, because of the quantity of the information and because of the technological level of the product. Because these products have so many features, they are - necessarily - complex. This means that, in spite of a well-designed user interface, much information remains to be transferred for each product. As has been explained in the previous chapter (and in Westendorp, 1998), the documentation has to be extensive because the products tend to be complex (and because of the wide range of users). Moreover, electronic and software engineers are still developing new products with many features and new features for existing products at a higher speed than ergonomists and user interface designers can cope with. Until a revolutionary user interface is found, extensive additional support for the user remains essential.

A large quantity of information implies detailed ordering and structuring of the information (table of contents, index, headers, footers, hyperlinks, etc.) and various systems that help users find the information they need. Until a revolutionary help design is found, a large quantity of information implies complex information. The presentation of the user instructions has to make this vast and complex information usable in an efficient and effective way. But it may also add to the complexity of the help system: for one user it may be useful that the desired information can found in several different ways, for another user this may be an additional complexity.

The way the information is presented can be split up into two elements: (1) the way to find the information (which could be considered as a part of the channel) and (2) the presentation in narrower sense: the graphic display. This analysis is based on the design process for on-line help systems that has been described and applied by Duffy, Palmer & Mehlenbacher (1992)[27]. It is clearly an analysis for on-line help, but most elements can be found in paper documentation as well. It seems useful to describe the analysis for on-line help, since this may be the most complex.

(1) The way to find the information consists of four elements: access, selection, navigation and application. A user first has to find out where to locate the help system. This is followed by how to reach the information that presents the solution to the problem. Next, the user may want to navigate to additional information. Finally, the user has to apply the information. Modern software may offer possibilities to perform the tasks that the user wanted information about, directly from within the help system. For selection and navigation, the designer may offer browsers (like: 'back' and 'forward', 'next' and 'previous', a 'home' button and a history of previously chosen items), bookmarks and other (hyperlinked) aids.

(2) The presentation consists of the following four elements: the choice of the media that will be used, the choice of the structure of the information, the choice of an overall graphic design and the choice of all the graphic details.
In on-line presentation of user help, the designer may choose a combination of several media: text, pictures, sound and animation (moving pictures, including video).
The structure of the information, as represented in the table of contents, may vary. For instance, it is possible to present the information for the main functions first and for the additional

---

[27] Duffy, Palmer & Mehlenbacher (1992) distinguish the following user tasks on which the design steps should be based: 1. Represent the problem. 2. Access the help system. 3. Select a topic. 4. Scan the information. 5. Obtain the required content. 6. Understand the information. 7. Navigate to other topics. 8. Apply the information.

functions later or to describe all functions following the ordering of the buttons or menus and menu items. On another level, the document designer may decide to present the descriptive information first or procedural information first. Indeed, the alphabetical index may be the only structure, as in the original documentation for Word Perfect for Macintosh. The graphic design encompasses elements like the choice of a visual metaphor, the size and layout of the screens, the use of (multiple) columns, the choice of picture style, etc. Finally, the designer of the instructions will select typefaces, icons, drawing styles, use of colors and many other elements from a large number of possibilities for each separate item.

One other important presentation choice may be made for instructional help in electronic products with screens or displays: the help may be context-sensitive or context-insensitive. Usually, a combination of the two systems is offered. In general, the context-sensitive help only presents descriptive information, no instructions (procedural information). Context-sensitive help may be presented under one button or presented with a help button on each window or even next to each separate item. In the first case, the cursor usually changes into a help cursor (like a question mark); if one selects an item with this cursor this item does not perform its function, but information about it is given. If a help button of a context-sensitive system is selected, only the information about this specific element is presented.

Designers of help systems will have to make a choice in the combinations of all these major and minor elements that have been discussed above. They have to design a concept and fill in the selected details. It will be clear that, because of the large number of design selections that can be made, many different concepts are possible, even if only the major elements are considered. The user may expect and indeed is confronted with a large number of possible conceptual presentations of the message.

3.1.5 Feedback

Feedback process

In the process of communication between designers of user assistance information and users, feedback has traditionally been a weak point. Authors of books on how to write technical documentation always stress the need for product tests, user assistance tests and user inquiries. However, practice is different: the management does not want to lose time in marketing the product and does not want to spend much money in after-sales activity.

But there are some signs of a turning of the tide, especially because of the Internet. Getting relevant feedback from the users has become easier (1) and getting large-scale feedback from the users has become more relevant (2).
(1) It is currently easier to get feedback than, for instance, ten years ago, since users consult help more and more often from an electronic database operated by the producer. Until now, these databases usually consist of Acrobat Reader files, which are just electronic versions of paper manuals. But there are a growing number of really interactive help systems that record the users' questions and attempts to find answers. Producers can also learn about the problems users may have from frequently asked questions, which are often also part of the producers' website. Once producers use this information and apply it in the new version of the product, we can speak of real feedback communication.
(2) Collecting and applying feedback has also become more relevant for producers, since the users communicate worldwide about the product, for instance in electronic user groups and discussion groups. As a result, product criticism is known worldwide in a very short time, and so are better products. Until recently, it was the single user complaining in the shop or on the phone, but now things have changed: users with complaints increasingly unite (virtually) and thus form a group that cannot easily be ignored. Moreover, new product versions (or: releases) are marketed so rapidly, that users have many occasions not to stay with the same manufacturer.

Product feedback

Of course products may give feedback by doing what they are supposed to do (the engine starts, the radio produces sound). Electronic products may also give (additional) visual or auditory feedback via displays, screens or speakers. Many of these products give error messages. These messages may just present the code of the message or describe it. But error messages may also contain instructions about how to solve the problem or immediately offer guidance to solve the problem (comparable with the interactive help). The system may also receive feedback from the user and learn from it;

the machine can then adapt to its user and automatically correct mistakes or present tips.

3.2. Conclusion

The overview of all sender, message, channel, receiver and feedback elements in the human - product interaction shows the complexity of the process (figure 3.7). This complexity should be realized when we consider conclusions that are drawn on the basis of studies concerning one (detail) element of the process. Especially for conclusions drawn for practical implications, the fact that the communication process in technical instruction for modern electronic products involves so many aspects may be a major limitation. Maybe that is why graphic designers sometimes complain that they cannot use the conclusions from science. Results concerning one design aspect cannot be applied in isolation from other design aspects.

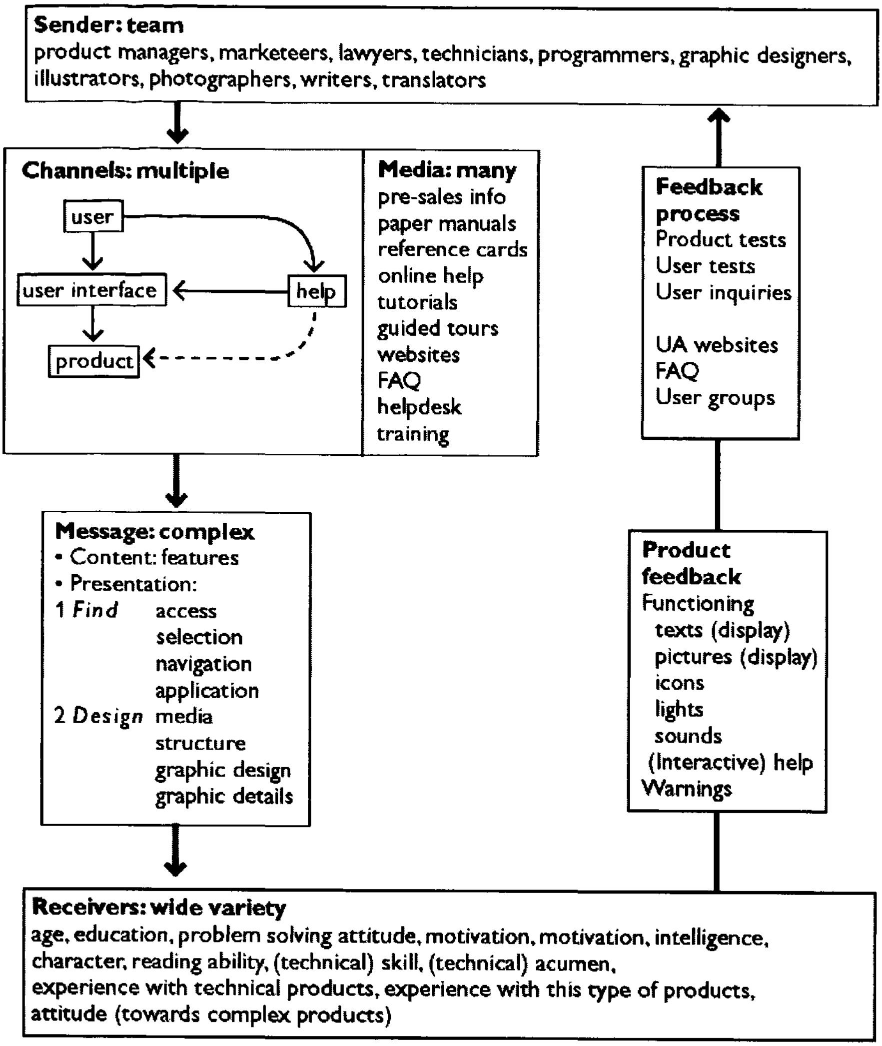


Figure 3.7.
User assistance information process.

CHAPTER 4

Human-machine interaction via user interface and user instructions

As we have seen in the previous chapter, the communication process between a designer of the instructions and the users is very complex, because so many actors and aspects are involved. The communication and interaction process between user and user interface is certainly not less complex, especially since the user has to deal with a complex product, a complex instruction, and switching back and forth between these two. These aspects of human-machine interaction also need to be described here, to be able to interpret conclusions about specific elements (such as the relative effectiveness of text or pictures) in the right context. In this chapter, the communication and interaction process between a user and a product will be described, focusing on the switching the user has to do between user interface and user instructions.

4.1. 'Instructions + tasks' model

Many theories and many models have been developed to describe the interaction between user and machine (for overviews, see Preece, 1994 and Coe, 1996), and various theories and models have been developed to describe the way the human mind works in a problem solving process, but it is difficult to find a model that specifically describes the interaction between user and product taking into account the user instructions as a separate source of information. The only exception may be the model that Jansen & Steehouder (1989) have proposed to describe the process of people that fill in forms. Van Hees (1994) has further developed this model to describe the interaction of humans with products and user manuals.

The models of Jansen & Steehouder and Van Hees are unique because they specifically take into account that someone who has to fill in a form or operate a product applying separate information such as a user manual, has to switch between this information and the form or the control panel[28]. Van der Meij (1998) discussed this problem only within the context of the minimal manual (§ 6.1.4). He discerns three separate sources of information (paper manual, screen and input device) and notes that users can comfortably deal with the separate handling of each of these sources, but that the joint processing is taxing. It is this coordination (Van der Meij also uses the terms 'harmonize' and 'synchronize') that makes instructions difficult to use. Sweller & Chandler (1994) called this a 'split-attention problem' and argued that quite a number of mistakes in the use of complex products originate in this split-attention of the users. Preece (1994, p. 311) notes that users of tools to help to accomplish something are engaged in two tasks: the primary task that leads one to use the tool and the secondary task of mastering enough of the tool to accomplish the primary task.

During this switching the user has to process and remember this information both from the form or product user interface and from the user instructions. When installing and operating a product, the user may have to switch frequently between user interface of the product and the user instructions and vice versa. The user has to switch during every detailed step of an instructional procedure and to check the feedback that the machine may present: read one step of an instruction, switch to the product, perform one action, notice the feedback from the machine and check this in the user manual. This may be a complex process, especially if the user manual has the pictures separate from the text and a mutual reference system (with numbers or letters); in this concept of a user manual, the overview picture is often on a foldout page (figure 4.1).

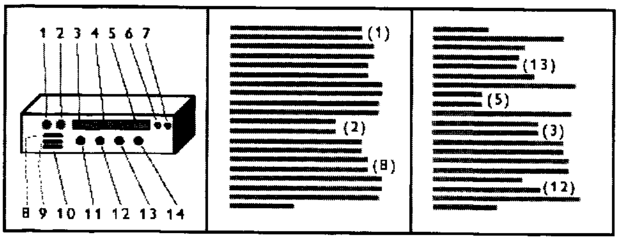


Figure 4.1. Well-known concept for a VCR manual: separate foldout page with illustration and referencing lines and numbers. In the text are references to these numbers.

In this example, users may have to make many switches to perform one single task. A user interacting with a VCR and a user manual based on the text in the booklet and numbers referring to an overview

[28] A situation that is comparable is when someone has to find his way through a city, consulting a map. In this situation too, the user has to switch regularly from the information to reality and vice versa.

drawing on a foldout page (a common concept for these kinds of user manuals) may have to perform the following actions and switchings for one single procedure (e.g., 'Set the volume to level 20 with the Volume control [12]') (table 4.1).

Table 4.1. Example of the switching that a user has to do when trying to operate a product, and consulting a user manual with an overview picture on a foldout page and a numbered reference system.

1 On the control panel	1. Notice a problem
2 In the user manual	2. Find the instruction
	3. Read and understand the instruction.
	4. Notice the number that refers to (the location of) a certain button that has to be pressed.
	5. Remember this specific step of this specific task and the number.
3 On the fold-out page	6. Find the number around or in the overview picture.
	7. Follow the line to the drawing of the button.
	8. Perceive and remember the shape and location of the button.
4 On the control panel	9. Find the location of the button.
	10. Recognize the button.
	11. Perform the task (by turning the button).
5 On the display	12. Find the display.
	13. Notice what it says.
	14. Remember what it says.
6 In the user manual	15. Relocate the task you were performing.
	16. Relocate the specific step.
	17. Check the feedback information (e.g., a separate small screendump-like picture of the display).

If the user forgot something or made a mistake (e.g., following the line from the number to the detail incorrectly), a part of the process may have to be repeated (meanwhile the user has to remember the instruction). This makes the process even more complex. The user may have to repeat this small-scale process for every single step in a procedure - and it is not uncommon to have procedures of seven steps each - and many procedures. Note that this only describes the switching for one simple procedure; the user may have to perform dozens of these procedures to install and operate a VCR and also has to process descriptive information (e.g., information that describes the functions or buttons).

This description is somewhat arbitrary, but it does indicate the importance to take into account the switching between product and user manual and within the user manual and vice versa, and to mention it in the model that describes the interaction process between user, user manual and user interface. This observation is supported by the many mistakes that seem to be made (e.g., in tests) during this switching process. Many researchers will recognize the situation in which a subject has read a certain instruction and then clearly does something totally different on the product's user interface than the instruction states. The subject may indeed seem to perform not only another step, but follow a completely different procedure (e.g., one that is described on another page). Often, such (kinds of) mistakes cannot be attributed to not understanding the information or misinterpreting the information. The conclusion must be that the subject has read and perhaps even understood the information, and while switching has either forgotten what to do and subsequently - while switching from user manual to product - remembers different information, or even makes up other information and tries to perform that[29]. The user may deduce a completely different operation principle - even from a different product[30].

4.2. The 'forms + instruction tasks' model

Jansen & Steehouder (1989) developed a model to describe the interaction-communication process of someone who has to fill in a form using additional information. Jansen & Steehouder discern three kinds of tasks:

- interpretation tasks
- functional tasks
- monitoring tasks

---

[29] Wagenaar, in particular, has published many articles about the limitations and 'fantasy' of the human mind. See for instance Wagenaar & Embrey (1990).

[30] Personal observation when executing tests for user manuals for Philips Car Stereo and for Philips VCR manuals. Several times a subject, who had to perform a task on the car stereo, read the correct instructions for that task and then did something totally different on the equipment. Later interviews showed that some subjects had followed a procedure for a comparable task on their own car stereo equipment.

The interpretation task includes the reading and understanding of the instructions (the reader must understand the information and infer the procedures). The functional tasks include the generation of data (e.g., a name, a date or an amount for filling in a form), the verification (e.g., checking to which category this data belongs) and the transformation (e.g., transferring a date into a code). The monitoring embraces the tasks that the reader has to perform to be able to switch between the interpretation tasks and the functional tasks. Jansen & Steehouder discern four monitoring tasks (figure 4.2):

1. orientation
   (develop a general idea of the paper form and of the task)
2. selection
   (deciding what steps have to be performed at a certain moment)
3. checking
   (checking whether the task performance has the desired result)
4. switching
   (stopping with the performance of interpretation tasks to start with functional tasks or vice versa).

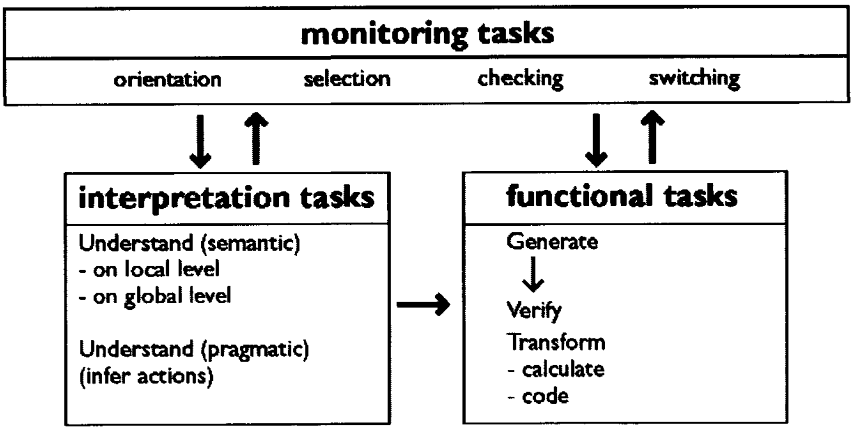


Figure 4.2. The Jansen & Steehouder model for the description of tasks that have to be performed when filling in a paper form and consulting separate information.

Jansen & Steehouder define switching as one of the monitoring tasks. They point out that 'switching' should not be confused with 'interpretation' or with 'selection', because these monitoring tasks concern monitoring within the same category (interpretation tasks or functional tasks), whereas switching is the moment when people go from one category of tasks to the other. This indicates that switching is another kind of monitoring task than the other three monitoring tasks and that the model should be slightly adapted to indicate this. Indeed, one could question whether this is a monitoring task at all.

In a test, Jansen & Steehouder found switching too early or too late to be a major problem. This does support the observations mentioned above that people may perform a task that is different from the one referred to in the instruction they have just read. Jansen & Steehouder did not try to find other types of problems caused by the switching itself. Perhaps switching is less of a problem for people who have to fill in forms than for those who have to operate a product using a user manual:

- the information may be very close to the place where people have to fill in the data;
- when filling in a form, people are reading and writing, so they are only concerned with language; when using a user manual to operate a product, people are reading and carrying out three-dimensional tasks, and this difference seems much bigger. They have to switch between language and a product.

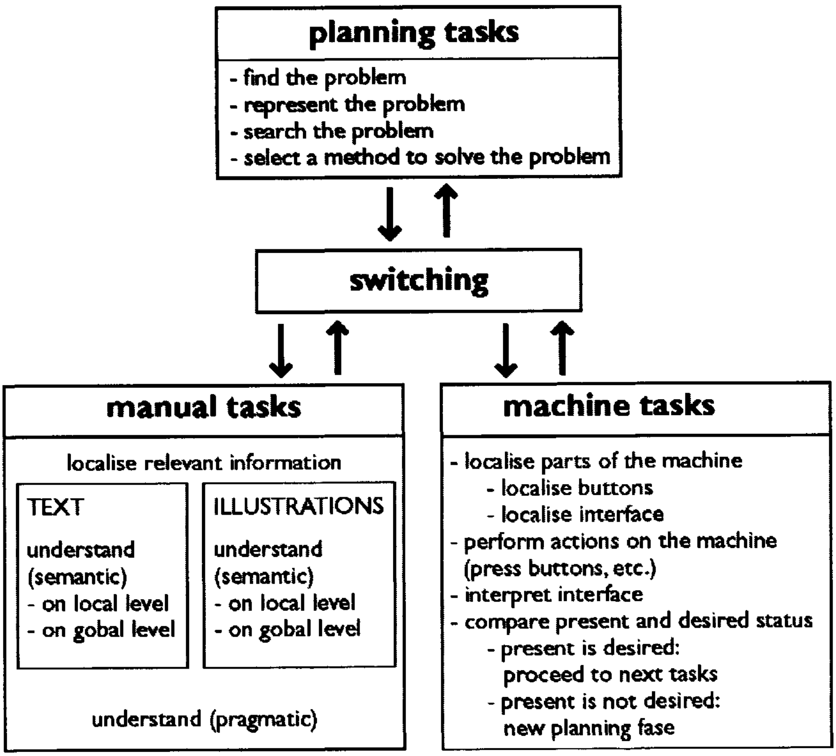
4.3. The 'product + manual tasks' model

Van Hees (1994) modified the Jansen & Steehouder model to describe the interaction between a user using a user manual and a product. She described interpretation tasks as 'manual tasks' (tasks that are performed by consulting the user manual) and 'machine tasks' (tasks that are to be performed on the machine). Referring to the problem-solving model by Newell & Simon and Hayes' vision on the problem solving process, she concluded that planning should be a separate element in the model. This planning would have four steps: find the problem, represent the problem, search the problem, select a problem solving method.

Van Hees further argued that checking, orientation and selection are in fact tasks that are done either when consulting the user manual or when interacting with the product. In her view, checking consists of the interpretation of a signal from the control panel of the machine and the checking of this feedback in the user manual. She therefore suggested putting these factors in the 'manual tasks' and in the 'machine tasks' categories respectively. She further argued that selection and orientation are not monitoring tasks when a user interacts with a machine and consults a user manual. In her view, these are both tasks that belong in the category 'machine tasks' as these tasks refer to localizing elements on the machine (control panel or elsewhere). On the basis of theories about information-seeking behavior she concluded that localizing the relevant information is a task that also has to be performed when consulting the user manual, and therefore should also be mentioned in the manual tasks module.

Van Hees noticed that there is no direct relation between interpretation tasks and functional tasks (the manual tasks and the machine tasks in her model), because there is always a switching phase. Therefore, she did not include the arrow between interpretation tasks and functional tasks as it was presented in the Jansen & Steehouder model. Van Hees also concluded that the switching tasks should not be limited to the switching between 'manual tasks', 'planning tasks' and 'machine tasks'. Switching also occurs between tasks within each of these three categories. Jansen & Steehouder did not account for these 'within module' switchings. Van Hees further added a component to account for the understanding of the information in pictures, apart from the understanding of information in text. She suggested that Jansen & Steehouder did not include a separate component for the information

Figure 4.3. The Van Hees 'manual-product' model.



in instructions, because pictures hardly occur in instructions for filling in forms.

Taking all these changes into account, Van Hees proposed the following model presented in figure 4.3.

4.4. Discussion of the 'product + manual tasks' model

(A) Planning

Van Hees rightly concluded that 'planning' is missing in the Jansen & Steehouder model. However, it is questionable whether 'planning' should be a separate component or module. The notion 'monitor tasks', as mentioned in the Jansen & Steehouder model, does incorporate a kind of process-control and planning could well be considered as one of the monitoring tasks. One could indeed also argue that these tasks would better be called 'planning tasks', but then the discussion only refers to the name. There is no need to follow that discussion any further. Moreover, it seems obvious that users of products, when consulting separate instructions, do have to perform certain monitoring tasks. They have to keep in mind where they are in the process of installing or using the machine, and it seems a bit forced to put all of the tasks for this process either in the 'machine tasks' module or in the 'manual tasks' module. One could also argue that planning tasks are either manual tasks or machine tasks or both. As a result of this, it seems better to have a 'monitoring tasks' module, which includes orientation, selection and checking and planning, but which does not include switching.

(B) Switching

Switching is an important element in the process of learning to install and use a product. It is difficult, however, to define exactly what happens in the user's mind during this switching. As has been argued above, switching should be an element in the user-machine interaction process, even if only to account for the fact that users have to store information in their short-term memory and may forget or - in reverse - imagine a (part of an) instruction when switching from the user manual to the machine or vice versa. The difficulty in defining what is or should be happening during the switching is expressed in Van Hees' model by the fact that she has not mentioned any task in this module.

Van Hees applied the same term 'switching' for both going from the instructions to the product (switching between modules) and for proceeding from one instruction to another (within a module). This seems undesirable, for two reasons. First, there seems to be a big difference between really stopping to pick up information and starting to act on the product (switching between modules) and proceeding from one step of information or one task to the next (switching within a module). Switching between modules means to stop reading, turn the eyes to the control panel of the product and start executing what one has read. Switching within a module means, for example, the mental step from localizing the information to understanding the information (in the 'manual tasks' module), or from localizing a button to actually pressing it (in the 'machine task' module) or from finding a problem to representing it (in the 'planning tasks' module). In fact, these within-module steps are not switching, but rather decision-making steps. Occasionally, proceeding from one phase to the next may imply real switching. For instance, to proceed from localizing the information in the user manual to the understanding of the information, the user may in between once again check on the control panel whether this is indeed the needed information. But this is in fact switching between modules.

Moreover, Jansen & Steehouder rightly remark that switching should not be confused with the pragmatic aspect of interpretation and that switching is different from selecting, because when switching the user actually stops reading and starts acting (or vice versa), whereas proceeding to selecting remains within the interpretation tasks. This is just one more reason to consider 'switching' as a separate element, not as being part of the 'monitoring tasks'.

(C) *Separation of text and pictures*

Van Hees included separate sub-modules for the processing of the information in text and in pictures respectively. However, the representation of the information should not be in this model at this level. First of all, for the instructions for installing and operating machines, more types of media are possible, notably sound and animations. Furthermore, it is still unclear whether text and pictures are being processed separately in the human mind. The discussion about the dual coding (as proposed by Paivio) or single coding (as supported by Pylyshyn) of textual and pictorial representation in the human mind has not reached a unanimously supported conclusion so far. We simply still do not know. Without clear motivation, Van Hees has chosen for Paivio's dual coding theory, although this has been under severe criticism (discussed later in this chapter in more detail). There are two more reasons why the representation should not be accounted for in separate elements in this model. First of all, the representation of the information (both internally and externally) can hardly be considered as a phase in the interaction process. Even if the representation would be processed separately for text and for pictures, it is not a step taken consciously or unconsciously in the interaction. Secondly, it is only sensible to present the form of representation, after the content of the information in the instructions has been specified (for example, whether it is declarative or procedural information or specified according to the taxonomy of types of information in instructions of Bieger & Glock, as will be discussed in chapter 6). It seems that the representation should be one level behind or deeper than the description in this model. In conclusion, it is better not to make a separation of various media in an interaction-communication model.

4.5. Repetitive 'interpretation - action process' model

The models that Jansen & Steehouder and later van Hees have proposed describe the repetitive interaction process between a user and a paper form or a machine; it is exactly that which makes these models unique. With the notes mentioned above, an adaptation of this model is represented in figure 4.4 to describe the repetitive interactions that a user has to perform on a machine. The main modules are 'interpretation tasks', 'application tasks' and 'monitoring tasks' and 'switching'. 'Manual tasks' seems too much a description of where the reader/user is looking, whereas Jansen & Steehouder rightly tried to describe the kind of tasks the reader has to perform at this moment. Therefore, their term 'monitoring tasks' is applied in this model and not Van Hees' term 'manual tasks'. The term 'functional tasks', that Jansen & Steehouder apply also seems not to describe what the user has to do on the machine. Because the user has to apply his or her knowledge, the term 'application tasks' is introduced for this module. 'Switching' is a separate module, but no content has been specified. In the 'monitoring tasks' module, 'planning' has been included as a separate element. In the 'interpretation tasks' module, Van Hees' element 'localization of the information' has been included, in combination with Jansen & Steehouders' 'semantic' and 'pragmatic understanding' elements. In the 'application tasks' the elements that Van Hees mentioned under 'machine tasks' have been copied with some minor changes (figure 4.4).

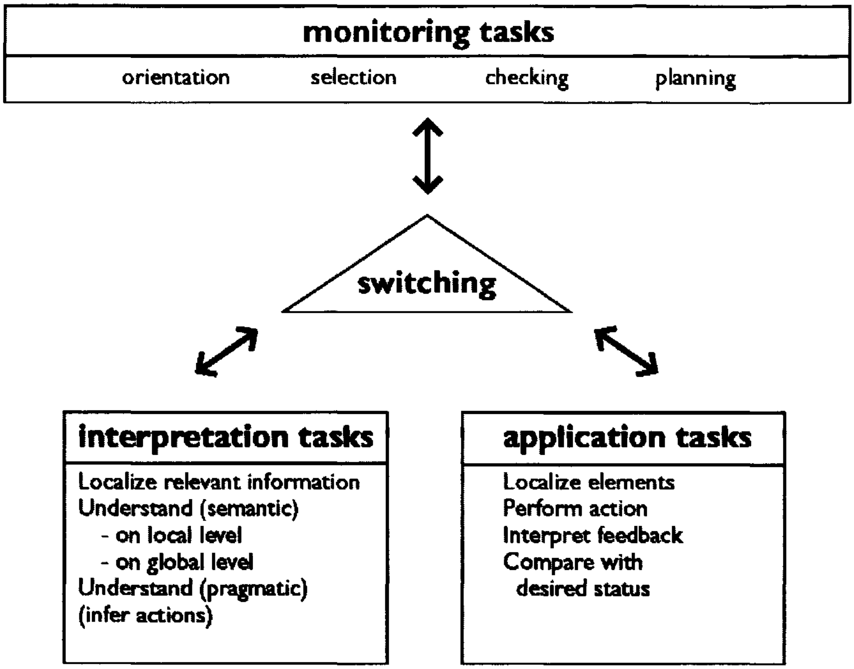


Figure 4.4. The adapted instruction-action model for the description of repetitive interaction between user, manual and machine.

This model emphasizes the fact that users have to switch between the user interface of the product and the user instructions. This switching between the user interface and user instructions (between interpretation tasks and application tasks) will be discussed further in the next chapter, where the psychological aspects involved will be studied in more detail. This will lead to (1) a discussion of perception of the user interface on the one hand and the user instructions on the other; (2) to the development of (abstract) mental models based on a combination of these two and (3) to the processing of various media in the human mind on the basis of the switching between user interface and user instructions.

Perception and cognition via user interface and user assistance

5.1. Hierarchical goal-orientation

In their interaction with modern technological products, human beings act goal-oriented. Although their goals may be quite vague and diverse when they are confronted with new products with many features, and although they may not even know all the features, this fits very well with theories that emphasize the goal-directed nature of behavior in general. Examples of such theories are Tolman's sign theory (Tolman, 1932) and Newell & Simon's theory about problem solving (1972). Tolman's basic idea is that organisms learn by pursuing signs to a goal, i.e., that learning is acquired through meaningful behavior. One could doubt, however, whether we are indeed always acting strictly goal-oriented when we interact with a technological product - or even that our behavior is always meaningful in these situations. Some functions may be discovered while we are trying to reach a certain goal, others by pure serendipity. We may be pursuing various goals and sub-goals at the same time and these may vary continuously while interacting with the user interface or communicating with the user instructions. Moreover, the ordering of goals and sub-goals is not a very active process, so it will be difficult to discover clearly organized hierarchies of goals and sub-goals in the minds of users of complex products. In general, though, the basic idea of the goal-oriented basis of learning seems appropriate in the human-machine situation. These theories specify that sequences of instructions have to be based on goals and sub-goals to be achieved, i.e., they stress the hierarchical nature of instructions[31]. Both Tolman and Newell & Simon stress this hierarchy of goals and sub-goals, described by Tolman as patterns of goals and sub-goals. The perception and further cognitive processing (integration) of the spatial-graphical user interface patterns and of the procedural user assistance patterns will be discussed in this chapter.

---

[31] These theories are opposed by behaviorist theories (Skinner), which stress the linearity of actions and reactions and therefore support a linear sequence of instructions.

5.2. Hierarchical information processing in human-machine interaction

If human beings act goal-oriented (to operate devices) and order their goals hierarchically, then they must also process the necessary information in a goal-oriented and hierarchical way - and this information must be presented in a hierarchical structure (Dixon, 1987a, 1987b, 1987c[32]). Taking this as a starting point, it is now time to describe this goal-oriented machine-interaction and information processing in some more detail. To indicate the line of thought within which this research has been done, in the following paragraphs, first a short overview of the most relevant aspects of human information processing is presented. Once this line of thought has been specified, some aspects that may directly concern the processing of various media (such as text and pictures) will be discussed in some more detail.

The interaction between user and machine involves many psychological aspects[33]. In the next paragraph, some of these are presented following the process from perception to action. The items will be further discussed in the following sub-sections of this chapter.

In order to let modern technological products perform the functions that they have, we have to solve problems. To solve problems in human-machine interaction, we need to perceive information, which is offered externally in two ways: by the design of the product and its user interface and by the presentation of the supporting information (advertisements, brochures, folders, but especially: on-line help, user manuals, tutorials and demos), and represented internally (based on memory). The external information may be presented in various media, including text, pictures (either still or animated), sound or a combination of various media. The internal information may be presented in various representations (e.g., in text, in schemas or in

[32] Dixon (1987[a,b,c]) argues that a reader of instructions constructs a hierarchically ordered mental plan on the basis of linear text. This is probably also true for visual instructions that are presented in a linear way.

[33] Kearsley (http://tip.psychology.org) presents a nice overview of relevant psychological theories, domains and concepts involved in his 'Theory into Practice database'.

pictures or perhaps amodal[34]), but research has not yet found an unequivocal way to describe the way internal information 'comes to mind' (this will be discussed in further detail in the next paragraphs). The information may be derived from schemata or from mental models (this will also be discussed in greater detail in the next paragraphs). On the basis of this information the user makes decisions. Moreover, the user preferably learns from his or her actions, so that future procedures can be performed easier. This learning does not only imply storage of information in memory, but also abstraction from the specific facts and a generalization to rules and mental models from which such rules can be deduced. The user has to make many decisions about which rule to apply.

To reach our goals we first have to perceive information and then to process this cognitively. In psychology it is common use to split up the information processing in the two main categories 'perception' and 'cognition', but it is also generally accepted that it is difficult to fully separate these two categories: perception is impossible without cognition and vice versa. Moreover, cognition concerns a wide variety of aspects related to information processing such as memory, learning, decision-making, problem solving and the ability to abstract or generalize and to deduct from these abstractions or generalizations.

5.3. Perception in human-machine interaction

5.3.1. Perception as pattern recognition

Perception is a fundamental characteristic of life. In order to continue living, every organism has to perceive its environment and the changes in this environment. To operate the technical products that we want to use, we first of all have to perceive all relevant signs: user interfaces and supporting help. Following Gregory (1995), perception is defined as 'the process of recognizing patterns from the outside world'. These may be visual patterns or auditive or any other sensory pattern that can be discerned, but in this study the focus is on visual perception.

An important - generally accepted - view taken here is that we do not perceive individual particles or pulses (like dots on a screen), but always complete patterns at the same time (Gregory, 1995; Hoffman, 1998; Dirken, 1999, p. 347-348). In interaction with devices we may

[34] 'Amodal' here means 'not in a specified medium'.

perceive patterns at various levels of detailing: a button, a complete front view, the complete product or the product including the surrounding area, but we always see a pattern. A visual pattern, for instance, may have size, place, value, texture, color, orientation and shape (Bertin, 1981 and 2001)[35] or any combination of these. We always see a pattern, just like we always hear a pattern (even if it only consists of one tone for a short distance)[36].

What does this mean for the interaction with devices via both user interfaces and user instructions? In chapters 3 and 4, it has been argued that users of complex technological products may perceive information via two different channels: via the user interface and via the user instructions. Chapter 4 highlighted the switching that users have to do between these two channels. Each of these two channels has its own content and presentation and these channels must be perceived as two different types of patterns.

The difference in presentation varies from completely separate and visually different (like a printed instruction manual that comes with a camera) to more or less the same way as the product is presented (in integrated on-line help). In software there is a tendency towards more integration of the presentation of the help into the actual program (Westendorp, Jansen & Punselie, 2000). Here it is of importance to notice that users perceive and further process cognitively two basically different types of information patterns when they try to operate a product and use help. We perceive a quite different type of pattern when we look at the user interface from the one we see when consulting the (procedural) user instructions.

(A) Perception of the user interface

We perceive the user interface of devices and software as a spatial-graphical pattern where spatial elements with a special function have a special form or color and brightness (or perhaps texture). The visual pattern is a graphical ordering (hierarchy) of elements indicating what elements are on the user interface, where these elements are located

[35] Bertin's graphic variables are mentioned here (see: Bertin, 1981 and Bertin, 2001). For three-dimensional elements, more visual variables can be discerned. Moreover, movement (animation) should be mentioned.

[36] Likewise, to describe feeling, tasting and smelling, a pattern may be a very useful term to describe the experience.

and what they look like (Bertin's graphic variables may be applied to describe this pattern). For the eye, the user interface is very much like a two-dimensional picture[37]. The user interface also presents the information in a hierarchical way: elements are ordered in a certain way, indicating their relationships.

The information in these patterns can be described using the terminology that Bieger & Glock (1984/5 and 1986) have devised (will be discussed in detail in chapter 6). The user interface of a device presents mainly 'inventory' information (what elements?) and 'spatial' information (where?), which may be supported with 'emphatic' information (such as warning signs).

(B) Perception of the user instructions

An important visual characteristic of procedural instructions is that these are a series of elements, often procedures consisting of numbered short sentences and/or simple pictures (only one or a few details, highly stylized). The combination of these procedures indicates the hierarchy, which is sometimes expressed in a table of contents. Whether it is a paper manual or an on-line help system, the user instructions can always be characterized as a linear procedural pattern of language (spoken or written text) or pictures (stills or animation) or a combination of these two, often hierarchically structured. This is especially obvious when the information is presented in a flowchart.

The help may contain (1) explanatory and (2) instructional information. The final goal of the help is always to instruct: to inform users what they have to do to switch on functions, connect elements, etc. To reach this goal the user may also need information which is not instruction, but rather more explanation[38].

(B1) The explanatory part of the user instructions is often presented in body-text or an overview picture with a reference to words in or near the picture to inform the user what the indicated elements mean.

---

[37] Two-dimensional, because our eyes cannot see depth. We can only feel depth. For visual perception, we have to construct it with our brains.

[38] Ummelen (1997) discerns two types of information in instructions: descriptive and procedural. This division is based on the distinction between declarative and procedural memory.

Explanatory information in a help system may consist of several of the information types that Bieger & Glock specified (definitions and examples will be presented in chapter 6): 'inventory', 'descriptive', 'spatial', and perhaps also 'qualifying' (modifying other information) and 'emphatic'. 'Qualifying' and 'emphatic' information are meta-level information: they arrange or specify other types of information. Explanatory information in the user instructions presents to the user basically the same types of information as the user interface plus descriptive information.

(B2) The instructional part of the user instructions is visually perceived as a (hierarchically) structured combination of linear patterns of words or pictures or a combination of these. The instructional information is basically a procedure, but it may consist of several types of information: 'operational' (which presents the actions), 'spatial', 'covariant' (such as feedback information), 'contextual' (presenting the end result of a series of actions), 'temporal' (order), and sometimes some 'qualifying' and 'emphatic' information.

In conclusion: the user interface is a spatial-graphical pattern from which the user has to deduce linear procedures. The user instructions, on the other hand, are a predominantly linear, procedural pattern with supporting explanatory information, on the basis of which the user has to recognize elements and find their location. Moreover, the hierarchical structures of the user interface on the one side and of the user instructions on the other are very different. To reach their goals and sub-goals, users have to construct procedural patterns (series of steps), much like the ones presented in the procedural part of the user instructions. To reach these goals, users may need supporting explanatory information (which has very different types of patterns) and apply the procedures on the user interface (which again has very different types of patterns). This difference in perception is likely to influence the way the information patterns are further processed; this will be discussed later.

5.3.2.Perception as a construction

A second important issue here is that perception is considered not to be a passive pick-up of information from the outside world. Vision is increasingly considered to be an active, although totally unconscious, construction of the outside world. Perception is not determined simply by the stimulus patterns; rather it is a dynamic searching for the best interpretation of the available information (Gregory, 1995). Hoffman[39] (1998), for instance, following Von Helmholtz and Marr, states that vision is a process of constructing images on the bases of perceptual data. These data may not only involve sensory signals, but also knowledge of the many characteristics of objects[40]. From this point of view, perception is to a certain degree the construction of reality[41] ('re-cognizing'). Our brain allows past experience and anticipation of the future to play a large part in augmenting sensory information, so that we do not perceive the world merely from the sensory information available at any given time, but rather we use this information to test hypotheses of what lies before us. Perception becomes a matter of suggesting and testing hypotheses (Gregory, 1995, p. 221). The senses do not give us a picture of the world directly; rather they provide evidence for the checking of hypotheses

---

[39] The idea that perception is to a large degree a matter of construction is generally accepted. Hoffman (1998) is cited here, because this is one of the more recent books on this subject and because it gained a lot of attention (see for instance the reviews at amazon.com). Moreover, Hoffman takes the idea of construction further and makes it more specific by presenting the rules that govern this construction process. According to Hoffman, vision is an intelligent process of active construction. 'What you see is, invariably, what your visual intelligence constructs'. Hoffman holds that we see the same things because we construct the same things and that we construct the same things because we use the same rules of construction (p. 74). Hoffman compares these innate rules of construction with Chomsky's rules for universal grammar. The mediaeval Islamic scholar Alhazen and the 19th century German scholar Von Helmholtz already described perception as a process of unconscious inference. Hoffman argues that the term 'inference' still indicates too much a conscious activity. Perceiving, according to him, is an active, but totally unconscious process.

[40] Hoffman (p. 64): 'The image at your eye is discrete. You can think of it as like a pointillist painting by Georges Seurat, composed entirely of tiny and separate daubs of color. Your image, like Seurat's painting, has neither curves nor surfaces. Logic dictates, then, that you are their source. You construct each curve and surface you see.'

[41] The philosophical discussion whether the outside world is a construction of the individual or a physical reality that is just perceived differently will not be pursued here. A nice short article about this point is the report of a discussion about this point between the Norwegian philosopher Arne Naess (who supports the theory that there is no outside world but only constructions of the minds) and the British philosopher Alec Ayer (who supports the view that there is one and the same physical reality). This discussion was organized by the International School of Philosophy in Amersfoort, The Netherlands, under the title 'The glass is on the table' (Elders, 1971).

about what lies before us. Indeed, we may say that the perception of an object is a hypothesis, suggested and tested by the sensory data. We need various cognitive aspects to perceive (e.g., memory, the ability to abstract and deduce). 'Our brain has largely overcome the limitations of our sensory organs' (Gregory, 1995).

What does this mean for the interaction with user interfaces and help systems?
It implies that the user interface will never be perceived as just a combination of arbitrary colors and shapes or elements on arbitrary spots: users will always make an immediate reference to (a comparison with) something perceived before, to their a priori information about user interfaces. They will automatically and immediately try to give meaning to what they see on the basis of what they already know. They know that what they see is a user interface, that elements have to be touched to perform functions, that there will probably be some kind of structuring of the elements, that visual codes such as colors or icons may give clues, etc. Users perceive a new user interface on the basis of existing user interfaces, especially by relating the newly perceived user interface with user interfaces of the same type of devices, or what they consider to be similar devices. Therefore, the user interface is not perceived as merely the spatial-graphical overview described above. Every user immediately perceives not just layout and color, but also hypothesized functionality. This perception may be incorrect and may vary widely for various users, but we cannot separate this hypothesizing from the spatial ordering and coloring. Paraphrasing Gregory, we could say that we do not perceive a user interface, but dots to which we immediately link a number of hypotheses concerning the functionality of a device and the way to operate the various functions. We do not see a user interface, but 're-cognize' it on the basis of our experiences with user interfaces that we have seen previously. Therefore, various users may perceive different user interfaces of one and the same device.

Users also perceive the 'help' function as help information because they expect some kind of prototypical help, with a certain content, design and structuring. The further the presentation deviates from this prototypical idea, the more difficult it will be to first of all recognize that it is the help. Users re-cognize the help on the basis of experiences with previously used types of help. Here too, various users may perceive different 'help interfaces' of one and the same 'help' for one and the same device. However, there may be some common aspects on the basis of which users re-cognize the user instructions. One general character of all user instructions is the hierarchical ordening. Users perceiving user instructions may see a linear pattern of words or pictures and immediately hypothesize or expect (and therefore perceive) a certain hierarchical ordering of the instructions[42].

In summary, when users interact with complex devices and use help, they 're-cognize', infer or construct (in Hoffman's terminology) two different patterns of information; one for the user interface and a different one for the user instructions. On the basis of previous experience, they 're-cognize' an image-like graphical layout (the user interface) from which they will try to infer rules to operate the device (trial and error). Also on the basis of previous experience, they 're-cognize' the linear series of words and pictures as the hierarchically structured help information. These two different perceptions are further processed cognitively when the users try to operate the devices: they will solve problems, make decisions, memorize elements and procedures, build a mental model and try to generalize rules (generalized procedures) and remember to apply these. This cognitive processing will be described in section 5.4, especially focusing on the separate hierarchical patterns for the user interface on the one hand and for the user assistance on the other hand.

Before that, one more remark. Users perceive these two different types of hierarchically ordered patterns of information because of their expectations concerning user interfaces and concerning additional help, that is: because of their a priori models of either the user interface and of the user instructions. Already at first sight, the user tries to fill in expectations concerning the content, structure, design and detailing of both the user interface and the additional help. Because of the wide variety of these elements of both interface and help, many misinterpretations start already at this point.

[42] Only a very simple instruction (e.g., just one arrow indicating a direction) may be an exception to the hierarchical ordening.

5.4. Cognition in human-machine interaction

In chapters 3 and 4 it has been pointed out that users who try to install and operate technological products while consulting external help, are involved in a very complex process (sender, message, medium, target group and feedback are all complex aspects) (chapter 3), and that they have to switch frequently between the product (control panel or user interface) and the help (series of steps in user manuals or on-line help systems) (chapter 4). Moreover, it has been pointed out that the user interface and the user assistance are perceived as two quite different types of presentations. The user interface is usually perceived as a spatial-graphical pattern: a layout of buttons, switches, handles and a display with at most some product graphics (icons, symbols and text) indicating the function of each specific part. The user interface information is mainly descriptive, informing the user about the specific function of a button, switch or handle. The layout of the buttons presents the user with spatial information (where to look for a button) and the organization (e.g., the combination of elements in menus or the color-coding of buttons on a control panel) may inform users about a hierarchical ordering of this information. The information in the user assistance, on the other hand, is perceived as a predominantly linear pattern, which presents procedural information. This information tells the user what to do. The user therefore switches back and forth between a graphical pattern with descriptive and spatial information and a linear pattern with procedural information. These two different patterns, with their distinctive types of information should be combined and further cognitively processed. This will be discussed below.

What aspects of cognition are involved in this process? Users of complex technological products will perceive information in order to process this into operating procedures. This involves many aspects of cognition. Users have to solve problems (5.4.1), make decisions (5.4.2), and they must learn from what they are doing (5.4.3), so that they do not have to consult the user manual for comparable actions over and over again. This learning involves two important aspects: memory and abstraction, especially concerning the various cognitive loads. Problem solving, decision making, memory and abstraction will be discussed below, focusing on the development, recognition and reproduction of patterns of actions needed to operate devices. After that, the efficiency of the processing of various media (text, pictures, animation) will then be discussed (in 5.5).

5.4.1. Problem solving

Users of modern technological products have many problems to solve. If these users would act completely rationally, they would probably proceed along the following steps. First, the user has to conceptualize the problem; this means he[43] has to form a psychological representation of the device or elements of the device. When the user thinks the problem is clear and understood he may have to do some deductive or inductive reasoning. Then he has to develop or select a problem solving strategy, execute this strategy and evaluate whether the strategy worked[44]. Problem solving skills are related to many other aspects of cognition like schema recognition (the ability to remember similar problems) and pattern recognition (the recognition of familiar problem elements). In this generally accepted description of the process, the user has to develop procedural action patterns (a problem solving strategy) on the basis of a mental representation. This representation is based on (1) the user interface of the device, (2) previous knowledge of devices that the user thinks are comparable, and (3) on the instructional information. During operation of the product the internal representation will continuously be adapted. That is why the 'problem space' continuously varies, not just because users set themselves new tasks, but also because they may do so on the basis of continuously developing representations. The more tasks they have performed successfully, the more they may develop a generalized procedural representation of a product, i.e., discover the abstract pattern of how to operate this specific product. This abstraction will be discussed later. Of course, this development of a mental representation of the abstract pattern of the procedures is usually not a conscious active learning of a hierarchical pattern. Especially when consumers are trying to install and operate a new device, their main goal will be to get it going and not to get an idea of the basic principles of the user interface. They start solving problems on a sub-goal level and can only develop a representation of the hierarchy and likeliness of the ways in which they have to solve problems after having successfully completed several of these.

---

[43] 'He' = 'he' or 'she'.

[44] This outline of the problem solving procedure is based on Coe (1996); the theory is based on Newell & Simon (1972).

As has been stated above, Newell & Simon stress that the first step in problem solving is the identification of hierarchically structured sub-goals (cf., Dixon, 1987a). If this is indeed the way the human mind works, it is all the more important that both the features in the product's user interface and the user instructions are ordered strictly hierarchically and that the hierarchies are also presented alike as much as possible. One could at this point hypothesize that it is more important for the user that the features and information are ordered strictly hierarchically and that all hierarchies are identical as much as possible than that the features can be operated in easier ways (shortcuts). This would be in accordance with the idea that not only perception, but all cognitive processing (problem solving, decision making, memorizing and learning) is based on one single process: recognizing patterns at various levels of abstraction. It must be immediately stressed here that these hypotheses are just ideas based on thinking, not based on empirical findings.

In conclusion, when solving the problems to install and operate devices, users - more or less subconsciously - develop a hierarchical pattern of ordered procedures, which are patterns on a lower level themselves (series of steps). These procedures are based on (1) the expectations that users have of the device, (2) a mental representation of the device, (3) the user interface and (4) the procedures presented in the user instructions. This mental representation continuously develops when solving new problems (i.e., activating new features) into more or less abstract representations of how to solve new problems on the same product.

5.4.2. Decision-making

When solving a problem, the user has to make decisions. Decision-making can be analyzed in various ways. Psychologists have proposed a division in binary decisions and complex decisions. Binary decisions are made when we have to choose between 'either this or that', between 'yes' and 'no'. Complex decisions are to be made when there is a choice between a wide variety of possibilities, each having its own importance. We may at a certain moment have to decide whether we choose to perform action number 1, which may have the results A, B or C (... n) or we could at that moment choose to perform action number 2, which may have the results C, D or E (... n) (results may overlap). For each decision that we make, some results may be desirable and others undesirable. This is, for instance, the case if we have to choose where to go during our holidays: each country has its advantages and disadvantages. This type of decision-making has to do with weighing the various results against each other.

What type of decisions do users make in human-machine interaction? It is obvious that users of modern technological products, when using information that comes with the product, make many binary decisions on a very elementary level. Procedures are presented in steps and the user has to decide whether or not to perform each of these steps. The interesting question concerning human-machine interaction is whether users also sometimes have to make complex decisions when installing or operating a product. At first sight this definitely seems to be the case. The user may have to decide to program the VCR at a certain moment, or set the clock first or first try to record some program. However, if we split up goals that users of modern products may have into sub-goals and tasks into sub-tasks to the most elementary level, we could also end up with real binary choices for these kinds of decisions: we either do this or we do that. In decision-making during the human-machine interaction process, choices just seem complex: analysis at the simplest level always leads to binary type decisions that are ordered hierarchically. In conclusion, when users try to configure and operate new products, while consulting the help, they make series of hierarchically ordered binary decisions to reach their goals.

Arthur & Passini (1992) discern two other ways of decision-making: users may adopt the optimizing model or the satisfying model. They argue that people who have to find their way in complex buildings, cities or other geographical locations, will often not try to reach the optimal decision ('the decision which leads to the best possible solution'). They will usually follow the 'satisficing' model: 'the decision that will lead to the solution that is good enough to solve the problem at hand'[45]. For decision-making in relation to performing tasks to install and operate technical products, the definitions should be extended, because the optimal decision not only reaches the

[45] The differentiation between optimizing choice and satisficing choice was first made by Herbert Simon in 1955. Simon stated that satisficing (i.e., making a choice that is good enough) is the most common decision strategy.

optimal result, but also does this in the most efficient way. Alternatively, the definition of the satisfying model should be extended to: 'the decision that will lead to the solution that is good enough to solve the problem at hand, either via the most efficient way or not'. For example, to find one's way in a city, the optimizing model implies not only choosing the best possible solution, but also in the most efficient way: reaching a specific subway station instead of reaching just any subway station, (which would be the satisficing choice) also means finding the shortest way to the specific subway station. An example in human-machine interaction would be that a user not only finds out how to print exactly what he wants, but also finds out how to reach this goal in the most efficient way (this usually means with the smallest number of buttons to press). In this example, a user who makes a satisficing choice, may decide to follow a process that will print what he wants, but perhaps not in the desired typeface. More common is the example that the user knows that a certain procedure is more time consuming, but will get the job done. In both these cases the user makes a satisficing choice.

In terms of the human-machine interaction the fact that users often decide according to the satisfying model has important consequences. Usually the procedures to operate features of a product are as much as possible consistent with each other. Consistency in hierarchy of user interface and in additional help design is in fact one of the most important prerequisites for good user interface design. As soon as the user recognizes 'the system' behind the programming functions of a device, he has developed a kind of abstract mental representation (this will be discussed further below). Now, if the system or device is consistent but the user is not (because he applies deviating procedures - making satisficing choices), this user will develop an incorrect (satisficing) device model for this (part of the) device or perhaps develop no device model at all - that is the case as long as the user thinks that each procedure is completely unique and arbitrary.

This conclusion implies that user interface designers might in some cases better design an interface that is in accordance with user actions based on the satisfying model, than a user interface that is based on the optimizing model of decision-making. For instance, an interface designer might develop a procedure in a user interface that lets the user perform more steps to reach a specific goal than would minimally be possible. This could be a better user interface if more users would intuitively choose this method - although they would all perform more actions than strictly necessary each time they perform the procedure.

An example may clarify this. On a telephone, the programming of telephone numbers under pre-selection keys can be done in five steps; the programming of the type of bell-sound is done in five completely different steps. It may therefore be better to let users always follow a procedure that has as many identical steps as possible, even though this would require more steps for each procedure, because that is a way the users may more easily recognize as 'the procedure that will get them there'. The shorter procedure reflects the optimizing model, and the longer procedure reflects the satisfying model. The alternative could, for instance, be to design an interface with 7 steps for each procedure, whereby the first three and the last three are identical for all actions; this may be less efficient for some actions, but more efficient concerning the overall learning time for the operation of all features. Overall, the sooner the user recognizes the pattern of steps that should be made to reach each sub-goal, the more intuitive the user interface is, even if this sometimes implies pressing more buttons. The same argumentation can be applied for the design of the help: consistently using patterns that are as much as possible identical in structure and design may be more efficient than occasionally using shorter instructions.

5.4.3. Learning

When operating devices, users solve many problems and make many decisions on a very detailed level: for each procedure and for every step of each procedure. Such decisions may only have validity for this one special procedure or step of a procedure of the human-machine interaction. Also, users learn from what they did: (A) they memorize functions and positions of the buttons on the control panel and series of actions that they have to perform and (B) they develop a more or less abstract representation of the kind of actions that they have to perform to make features work. The user is probably quite unaware of these aspects of implicit learning[46], but somehow realizes that it could be useful to remember the procedure, because he may have to perform it again - preferably as quickly as possible without consulting the user

[46] As opposed to explicit learning.

instructions again. Furthermore, this user may - to a certain degree - also realize that it may be wise to learn to abstract from procedures, i.e., understand the rationale behind them so that he can more or less predict future actions. Both these aspects of learning, memory and abstraction, will be discussed in further detail and relevance for the design of user interface and user instructions.

(A) Memory

If we identify the human-machine interaction as a situation in which users perceive two different external representations of information (the graphical-spatial user interface and the linear-procedural help), two influential theories concerning memory should be discussed: Anderson's ACT theory and the Soar theory (Newell, Johnson-Laird). Both these theories are goal-based production systems and directed at real-time performance and learning; both theories aim at psychological validity (Aasman, 1995, p. 12).

In his Adaptive Control of Thought (ACT[47]) theory, Anderson (1983, 1993) contends that all knowledge begins as declarative information ('which refers to facts we know'). In his view, procedural knowledge ('production rules' or 'productions'), 'which refers to skills we know how to perform', is learned by applying factual knowledge, which has been forwarded from the working memory (short term memory). The ACT theory also postulates a propositional ('textual') basis for memory (Kearsley, 1998; Markman, 1999). In terms of human-machine interaction, Anderson's theory implies that all knowledge about a device is first of all descriptive 'how-it-works' knowledge in propositional form. On the basis of this understanding, users develop production rules to operate the devices.

To account for the memory functions in human-machine interaction with users consulting user instructions, the ACT model is useful because it specifies the differences between declarative knowledge and procedural knowledge. In user instructions, both declarative information and procedural information may be presented (Ummelen, 1997). The user interface usually only presents declarative information, the user instructions may give descriptive information ('what is this?'), but must always give procedural information ('how do I...?').

[47] 'ACT' is used here to represent all variations of this theory, including ACT* and ACT-R.

The Soar theory[48] postulates that memory is unitary and procedural. It is unitary because it presupposes that there is only one type of memory; for instance, there is no distinction between procedural and declarative memory. Memory is considered to be procedural, because all knowledge is goal-directed and to reach goals, human beings are supposed to develop procedures. The Soar theory postulates that human cognition can be described in the form of a production system ('heuristic searching in problem spaces') and therefore all information is stored in procedural form (Newell, 1990; Kearsley, 1998).

The Soar theory is useful here, because it focuses on the users' ultimate goal: to obtain procedures to install or operate a device. They may learn these procedures in different ways, by trial and error interaction on the user interface or by consulting the help (and all thinkable combinations of these). In the following it will be argued that this may lead to different types of mental representations. When users try to operate the product by trial and error, they focus on the spatial-graphical display. Such users may build up a mental representation especially based on declarative information. Users who decidedly and strictly follow the instructions, especially if they focus on the procedures, may develop a mental representation that is based on linear procedures and abstractions of these procedures.

Because learning the procedures is the final goal, all users store predominantly procedural information patterns in their memories. If they have to perform a series of tasks on the same device, they perform various related procedural patterns and store these in their memories. On the basis of these networks of procedural patterns users try to abstract more generally applicable rules to be able to predict future procedures to operate the device. Aspects of these abstractions into some kind of model that allows predictions will be discussed in the next paragraph.

---

[48] 'SOAR' means 'State Operator And Result', but the term is no longer used as an acronym. It is nowadays used as a term that stands for an architecture of human cognition expressed in the form of a production system, and therefore is now spelled as Soar. The Soar theory builds upon earlier cognitive models such as GPS (General Problem Solver) and GOMS (Goals, Operations, Methods and Selection rules) (Preece, 1994, p. 418). These theories are strongly based on the 'human processor' metaphor, the comparison of the human mind with a computer processor. This reflects why these theories are all unitary and procedural.

Organisms with brains do not only perform (a series of) actions to fulfill their immediate goals, neither do they just remember the specific reactions to specific situations. They are also able to recognize, discover or develop more abstract patterns on the basis of repetitive perceptions and series of actions (and their reactions in the physical world). This implies that they develop some kind of categorical mental representation of the world around them. 'By encoding our experiences into an organized system of categories, we are able to recognize significant commonalities in different experiences. A category system allows us to derive further information about an object that has been assigned to a category.' (Glass & Holyoak, 1986, p. 149). This (visual) categorization must be true for all organisms with brains, not just humans. A gazelle, for instance, does not learn to be afraid of one specific lion with its associated characteristics (visual, auditive, smell). It will have to learn to run away from any lion, whatever its color, size, silhouette, motion, sound and smell patterns. To survive, all animals must be able to abstract the perceptual patterns of their predators and preys. From this we may conclude that all animals have the ability to develop abstract mental representations (categories) from which they infer actions for specific situations[49]. We may therefore conclude that not only specific series of acts, but also abstract representations are stored in memory. This ability to recognize, store and recall abstract patterns (regularities and principles) is, of course, highly relevant for users of complex devices; they perform many actions on the basis of such mental representations.

Discovering regularities or principles on which a device is based will give users the ability to predict how to perform related actions and therefore make comparable future actions easier to perform. In order to discover such regularities or principles, users have to abstract from the specific steps in the series of actions. They have to recognize or discover a more abstract pattern on the basis of specific patterns: the similarity in the various series of actions that they have performed to activate features of a product.

---

[49] But the representation of such abstractions can probably only be propositional ('in linguistic form') for human beings. The idea of categorization in animals was explicitly raised and tested by Herrnstein. His famous experiments with pigeons have had many follow-up studies, most with pigeons. Herrnstein noted in 1984: 'while language depends on categorization, categorization does not depend on language.' Recent research indicates that pigeons are also capable of categorization that is not based on visual simularities, but on functional simularities (Urcuioli, 2001)

How such more abstract knowledge - consisting of regularities or principles - is mentally represented has been the subject of philosophical and psychological discussion for hundreds of years. Arnheim (1969, pp. 97-115) discusses this subject concerning the internal representation of images and the abstraction ('particular' and 'generic' images). He cites Aristotle, who already discussed the problem: 'Thinking is necessarily concerned with generalities. How, then, can it be based on individual memory images?' (p. 98).

Johnson-Laird (1988) summarized this 'traditional puzzle in philosophy' as follows: 'How can an image, or a diagram in a geometric proof, stand for many different things?' He then proceeds by citing Immanuel Kant who postulated in his *Critique of pure reason* that such abstractions of perceived patterns are best defined as schemata:

> 'In truth, it is not images of objects, but schemata, which lie at the foundation of our pure sensuous conceptions. No image could ever be adequate to our conception of triangles in general. For the generalness of the conception it never could attain to, as this includes under itself all triangles, whether right-angled, acute-angled, etc. whilst the image would always be limited to a single part of this sphere' (p. 101).

Johnson-Laird disputes this abstract, static representation and continues to cite Wittgenstein, who believed that specific objects might be prototypical, such as a schematic leaf, or a sample of green. Following David Hume's *Treatise on human nature*, Johnson-Laird concludes that the brain may indeed take a specific sample of reality as a prototype and then develop it on the basis of further experiences and perceptions. This continuous development of a prototypical sample of reality is the main reason for Johnson-Laird to indicate the internal representations of abstract patterns as mental models. Mental models are actively constructed and provisional representations of reality, and they can be revised in the light of subsequent information.

How this knowledge is stored and processed in the brain and which representation is most adequate, still remains an empirically undecided problem (Molitor, Ballstaedt & Mandl, 1989). Researchers have proposed a wide variety of descriptions of mental models. They may, for instance, be described as 'theories', 'rules', 'mechanisms', 'how-it-works knowledge', 'some kind of understanding of how a device works', 'structures', and the medium may be either images, schemes, propositions or flowcharts. In general, ergonomists specializing in human-computer interaction (Norman; Kieras & Bovair; DeKleer & Brown and others) often use the term 'device model' instead of mental model. They tend to describe a mental model or device model as 'how-it-works knowledge' or as 'some kind of understanding of how a device works' and an 'image'. Most definitions or descriptions of mental models or device models take the view that the understanding or description of how a device works is central. Some examples may support this conclusion (relevant aspects have been *italicized* here).

'A mental model is *a basis for explanation and understanding* translations of external process into words'.
Craik[50] (1943, cited in Rutherford & Wilson, 1991, p. 41)

'In a mental model, reality is represented in an *analogous, predominantly imaginative form.*'
Steiner (1988, cited in Molitor, Ballstaedt & Mandl (1989, p. 10)

'By mental model is meant some kind of *understanding of how the device works in terms of its internal structure and processes*'.
Rasmussen (1991, p. 42)

'*Contrary to the how-to-do-it knowledge a mental model represents the how-it-works knowledge*. A mental model is a representation of the internal structures and processes of a system at some uniform level of *description*. From a mental model any operating procedure for a device may supposedly be inferred by general reasoning processes. Therefore acquiring a mental model from which all relevant how-to-do-it knowledge can be derived may be overall more parsimonious than memorizing a collection of operating procedures'.
Schmalhofer & Kühn (1990, p. 337-338)

---

[50] Craik is supposed to have coined the term 'mental model' in 1943 (according to Philip Johnson-Laird, cited in Rutherford & Wilson, 1991).

'In previous research (Norman, 1983) it has been demonstrated that the users' mental model (i.e. *how-it-works knowledge*) is very important for *understanding* their interaction with the system. Some experimental studies have shown the advantages of explicitly instructing a mental model (Kieras & Bovair, 1984). Also, it has been demonstrated that *how-it-works instructions can be more effective than how-to-do-it instructions* (Schmalhofer, 1987), which may be due to how-it-works instructions being more general (Catrambone, 1988)'.
Schmalhofer & Kühn (1990, p. 358)

'Mental models are the mechanisms whereby humans are able to generate *descriptions of system purpose and form, explanations of system functioning* and observed system states'.
Rouse & Morris (1986, p. 349)

'Mental models are the *mental structures we use to understand* systems and solve problems arising from the way systems work'.
Winn & Snyder (1996, p. 112)

'Mental models are conceptual models *of how things work,* how things happen and how people behave'.
Norman (1983, p. 7)

'People often develop mental models, a kind of simple theory, of *how certain devices work* and use this model as an operational guideline'.
Keren (1996, p. 7)

'The *explanation of the functioning* is inferred bottom-up from a representation of the structure of the device in terms of the topology of component connections together with component properties'.
De Kleer & Brown (1983, p. 155)

'By 'mental model' is meant some kind of *understanding of how the device works* in terms of its internal structure'.
Kieras & Bovair (1984, p. 255)

Kieras & Bovair noticed the existence of two strong opposing intuitions. They found that psychologists were generally convinced that having a device model would be of great value. In technological industries, Kieras & Bovair noticed, the dominating opinion is that device models are unnecessary, and that instructional material should focus on how to get the job done and should not contain any how-it-works knowledge. Kieras & Bovair did some tests, the results of which indicated that providing a device model can result in faster learning and better retention of operating procedures[51].

In summary, the most common opinion amongst psychologists specializing in research of human-machine interaction seems to be that a mental model or a device model is 'some kind of understanding of how a device works' and that it is descriptive in nature. It is questionable, however, whether a device model indeed is some kind of understanding of how a device works. Many users of complex technological devices are not at all interested in how the device works; they are interested in getting the device to work for them. This means that they are primarily and perhaps only interested in knowing the procedures that they have to perform to install and use the device with all or some of its features. Therefore, one could argue that a device model is for most users rather more an abstraction of a series of procedures than a kind of more abstract descriptive idea of how it works. As has been argued before, these procedures are not just the low-level procedures for the operation of specific features, but also the more abstract patterns of procedures that users try to recognize: the patterns of the overall procedure that they have to apply to operate the various features of the device with variations for each separate feature.

It is argued here and below that, instead of a descriptive how-it-works device model, users rather build up a 'Procedure Model', which is an abstract idea of the pattern of the procedures that one has to apply to operate the features of the device. Such a procedure model consists of abstract procedural patterns. A procedure model is not an idea of how it works, but rather more a hierarchical collection of abstract procedures for the operation of the various functions of that device

[51] In this test two groups of subjects were compared, one group read a how-it-works description of the device and one group did not get this description. It seems obvious that the subjects who did get extra information (and maybe also extra time, Kieras & Bovair do not give information about this) performed better than subjects who did not get the extra information (and extra time?).

plus some idea of the topology of the functional elements of the user interface (the spatial information). The development of such a procedure model on the basis of repetitive actions with the same pattern of actions is shown in table 5.1. The ' and " indicate steps in a procedure that are comparable to the steps of the first procedure, but not necessarily exactly identical. For instance, the user has to press buttons from the same group of buttons (e.g., selecting from the buttons with the numbers 0-9), but not exactly the same button. Specific examples are presented further below.

Table 5.1. Development of a procedure model.
Uppercase indicates specific; lowercase indicates abstracted

	step	step	step	step
procedure 1	A	B	C	D
procedure 2	A'	B'	C'	D'
procedure 3	A"	B"	C"	D"
(...)				
procedure model	[a]	[b]	[c]	[d]

On the basis of this procedure model, the user may predict and try out other programming functions of this device. The user deduces new procedures on the basis of the abstraction of the ones he has applied successfully. This procedure model is much more a linear pattern than a 'how-it-works' device model or mental model, consisting of abstract procedures plus a topology of the user interface elements - which is not an abstraction. It is the abstract pattern of the serial actions that forms the mental representation for operating technical products.

The procedure model is not so much based on 'how-it-works' knowledge, but rather consists of 'how-do-I-operate-this-machine' knowledge. Users are not interested in the way the machine works, but much more in what they have to do to operate the various functions. If the various functions of such a machine are indeed operated via a more or less generalizable series of actions (parts of the series being identical), then the user may develop a more generalized pattern of operations. This means that the user does not develop a more generalized idea of the user interface (the topology), but develops a procedure model on the basis of the user instructions. Only if a user is learning to operate a product purely by trial and error, without consulting the user instructions, could one claim that the user builds up a procedure model on the basis of re-constructing the operating procedures via the user interface information (including feedback from the machine).

Such a procedure model is, just like a mental model, dynamic in the sense that the user is actively constructing it and that it may change continuously on the basis of new experiences. A user can only construct such a procedure model if the features of the device are indeed activated according to hierarchically ordered procedures. This is the case for most modern technological devices and software, although the abstract procedure may not always be easy to discover. Some examples may illustrate this argumentation that users of complex technological products with hierarchically ordered sets of features rather build up a procedure model than a device model or a mental model of the device. These examples also show the varying abstractness of the procedure model: example 1 (table 5.2) shows a procedure with one step that varies (select number on a VCR); and example 2 (table 5.3) shows a procedure with two steps that varied (programming sounds on a telephone).

Table 5.2. Procedure model with one variable.
Example 1. VCR: programming stations under pre-set buttons.
In this example only step three varies; the procedure model has one level of abstraction. Italics indicate abstraction.

	step 1	step 2	step 3	step 4
procedure 1	PROGRAM	SEARCH	1	PROGRAM
procedure 2	PROGRAM	SEARCH	2	PROGRAM
procedure 3	PROGRAM	SEARCH	3	PROGRAM
(...)				
procedure model	[program]	[search]	*[number]*	[program]

In this example, the procedure model that the user constructs will be: 'To program stations on this VCR, I always have to press: 'program', then 'search', then the number I want, then 'program'. This is a generalized procedure, on the basis of which the user may predict how to program the next station under the desired number.

Table 5.3. Procedure model with two variables.
Example 2. Programming sounds on a telephone: 1 loudness, 2 tone, 3 melody
This is a two-level abstraction: both selection of the type of sound and the selection of the sound intensity or variant vary. Italics indicate abstraction.

	step 1	step 2	step 3	step 4
procedure 1	PROGRAM	SELECT LOUDNESS	SELECT 1-5	PROGRAM
procedure 2	PROGRAM	SELECT TONE	SELECT 1-3	PROGRAM
procedure 3	PROGRAM	SELECT MELODY	SELECT 1-7	PROGRAM
(...)				
procedure model	[program]	*[select sound variation]*	*[select number]*	[program]

On the basis of repetition of these actions, the user may discover a pattern and therefore learn:
'If I want to program something on this telephone, I always have to start and end with pressing the programming key; in between I always first have to choose the function and then the variation of that function'. So the abstract procedure for programming this device is: [programming - selection - intensity/variant - programming].

The procedure model for activating the various features of this telephone may be further developed for other programming functions and thus lead to an even more abstract procedure (three or more levels of abstraction). For instance, the user may learn that the procedure for programming telephone numbers under pre-set buttons on this telephone system involves the following procedure:
[program] - [enter number] - [select pre-set button] - [program]
On the basis of this, users may take their idea of how this specific device should be programmed to a more abstract level, indicating: programming on this device always implies beginning and ending the procedure with pressing the 'program' button and always selecting from menus in between. If these users have this same experience with some other programmable devices, they may quickly develop the overall abstract procedure model that for programming devices one always has to begin and end with pressing the 'program' button - which, unfortunately, is not true[52].

[52] If this reasoning is correct, it may have an important implication for designers of user interfaces. If users indeed build up an abstract, linear 'procedure model' and base their predictions for future actions on this model, it may be more important that all features are activated and de-activated according to the same action pattern than that there may be shortcuts. As has been argued before, this is even truer if users make more decisions based on satisficing choices.

The procedure model, as described above, is an abstract pattern of series of steps that the user has to perform to operate specific features of devices. This model seems to deviate from the general opinion of psychologists specializing in human-machine interaction about mental models, but of course many researchers have pointed to the 'how-to-do-it' type of mental models. Young (1983), for instance, has argued that there are many types of mental models, the most important of which are the 'structural model' and the 'functional model'. A structural model assumes that the user has internalized the structure of how the device or system works in memory. The structural model represents how-it-works knowledge. A functional model assumes that the user has internalized procedural knowledge about how to use the device or system. The functional model represents the how-to-use it knowledge.

> 'Functional models develop from past knowledge of a similar domain and not - like structural models - from a model of how the device works. Structural models can answer unexpected questions and make predictions; functional models are structured around a set of tasks. Functional models are context-dependent; structural models are largely context-free.' (Young, 1983, p. 38).

There is an important difference between the procedure model proposed here, and the functional model as described by Young. The procedure model is the abstraction of linear patterns of procedures which does allow predictions for future actions on the same device, and possibly also on other devices that the user assumes to be similar. Youngs' functional model is defined as context-dependent and appears not to allow predictions for future actions. In fact, one could question whether users indeed ever build up more than a procedure model of devices. Perhaps the understanding of a VCR will for many users never be more than an internal representation of a control panel (with a graphical layout, a topology of buttons and other elements) plus an abstracted pattern of series of procedures. For the mental representation of a complex technological product, the procedures in the user instructions may be more important than the elements on the display.

5.5. Representation media

Now that the human-machine interaction has been described as a hierarchically ordered set of linear procedural processes of either direct input (first time use of a specific procedure) or on the basis of recognition and interpretation of procedural patterns in varying degrees of abstractness, it is time to discuss the medium in which such patterns may be stored in the brain.

Before discussing the various media in which the instructional information may be presented, it must be recalled that the user receives instructional information in two very different ways. On the one hand, users perceive the display, control panel or user interface (in the limited sense: without the additional information presented in user instructions). This information is presented in a spatial-graphical way; it often looks like a three-dimensional picture. Users may try out the buttons and other controls, but these themselves often do not present procedural instructions. On the other hand, users who consult the help system do perceive information in a linear procedural way. As has been argued above and in other publications (Westendorp, Jansen, Punselie, 2000), users may well develop different mental models of the same device on the basis of these two so very different types of information (figure 5.1). These mental models may vary from spatial-pictorially oriented (if the user has tried out the device purely on the basis of the user interface) to linear-procedurally oriented (if a user has consistently used the additional procedural help information). As has been argued above, the user develops one single procedure model that has both spatial-graphical information (the topology of the elements) and some form of more abstract-procedural information.

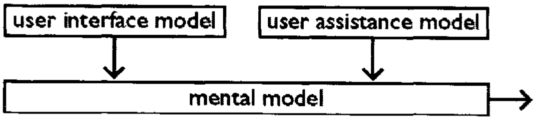


Figure 5.1. Development of a mental model of a device on the basis of graphical-spatial user interface information in combination with linear-procedural help information (also see figure 3.6).

This brings us finally to the question whether we perceive, store and remember pictures and texts separately as pictures and texts or rather in one, perhaps amodal, way. This discussion has had a long history, but no uniform conclusions can be drawn yet. For the instructions in the human-machine interaction, this discussion is interesting on two levels: on the level of the direct, concrete representation of the text or pictures that someone may just have seen, or on a more abstract level (in memory).

Following Paivio (1983), many researchers assumed that the human brain has two separate processing systems: a verbal system, specialized in processing linguistic information and an imaginal system, specialized in processing visual information (Molitor, Ballstaedt & Mandl, 1989). According to these scientists, a text is predominantly processed and stored in a verbal system, and pictures are processed and stored in a visual system.

The dual coding theory is in accordance with the 'mind's eye' hypothesis (Tabachneck-Schijf & Simon, 1996, p. 30), which states that, just as external representation has various modes (graphs, diagrams, words, equations, etc.), so internal representation has corresponding modes; and if something is learned from an external representation in one mode, it will accordingly be stored in the corresponding internal mode, and will again be in this mode when retrieved from memory. Thus, a sentence read will be stored internally in some kind of linguistic representation and a drawing will be stored in some kind of pictorial representation. In this view, each mode of representation has its own data structures and operators; pictorial operators cannot work on verbal data and vice versa. The 'mind's eye' hypothesis does not require that pictures or sentences in the head be exactly the same as pictures or sentences viewed.

This dual coding view has met strong opposition by those who assume that text and pictures are perceived modality-specific, but are processed further by one uniform processing system. These scientists (with Pylyshyn as the most prominent representative) assume that the perception of text and pictures involves two different modalities, but the further cognitive processing involves only one system. They state that there are no different parts of the brain for processing these different presentations of information. This view does recall the discussion above of the ACT and the Soar theories: the latter theory states that the human mind functions in a unitary and procedural way, following one single process. Researchers following the Soar theory are likely to assume that all information, perceived in whatever medium, is processed in one unitary way, just like computers process all information basically in one unitary way, whether the information consists of texts, realistic pictures, movies, tables, flowcharts or other presentation media.

In the context of this study it is relevant to notice some more supporters of the theory of one single, amodal internal representation of different media presentations. Amongst these are Chase & Clark (1972), Frederiksen (1975) and Stone et al. (1981) and Stone & Glock (1981). This last reference is an article that concerns the study of presenting various types of information in instructions for building a model loading cart in either text or pictures. This study was later continued by Bieger & Glock (1984/5, 1986) and served as the basis for the hypotheses tested in this study (chapters 7, 8 and 9) (the types of information will be introduced in chapter 6).

Many researchers take intermediate positions (e.g., Kieras and Kosslyn), supposing that we do have a separate imaginal system at a central level of processing, but that in the end all information is processed in propositional form. It remains somewhat unclear what this central level is exactly, but one could imagine that information is consistent in representation mode during perception and in short-term memory (STM) and transforms when the information passes on into the long-term memory (LTM). However, there is little argumentation for the thesis that the internal presentation changes from one medium to another during the transfer from STM to LTM. Nonetheless, there may be more arguments, but not much more proof for any of the other positions taken concerning the media of internal representation. This is in fact comparable to the situation concerning the relative value of the various external media for instructing users, or any learner or problem solver for that matter[53].

Molitor, Ballstaedt & Mandl (1989, p. 8) - on which this overview is based - present a clarifying diagram to express the three possible approaches (figure 5.2).

As has been stated above, in the human-machine interaction, users repetitively have to switch between user interface (display) and user instructions. This means that they have to switch between the spatial-graphical display that is comparable to a picture and instructions that

[53] '(...) that little is known about the cognitive value of any graphical representations, be they good old-fashioned (e.g., diagrams) or more advanced (e.g., animations, multimedia, virtual reality)' (Scaife & Rogers, 1996).
'The literature is overflowing with work investigating the facilitative effect of pictures on text comprehension. And yet, no one has a clear idea of the cognitive processes underlying these effects' (Clark, 1983, p. 446).

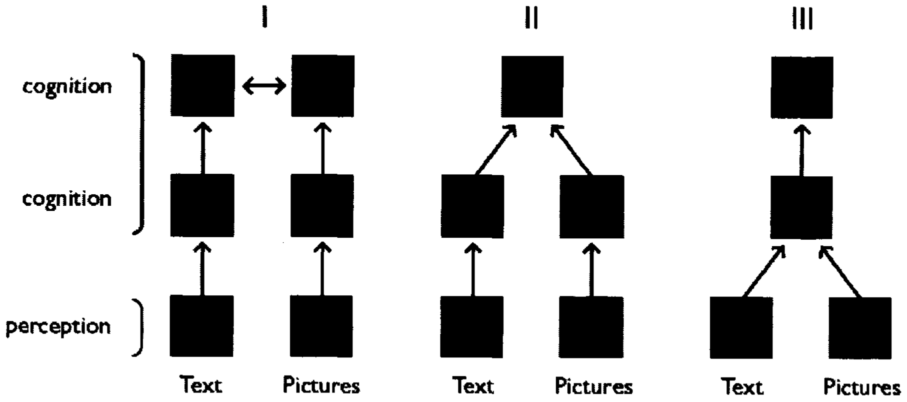


Figure 5.2. Do we process text and pictures separately, as Paivio contended (expressed in I), or rather with one processing system, as expressed by Pylyshyn (expressed in III) or do we have some kind of two-stage processing system for text and pictures (expressed in II, a position that Kieras and Kosslyn seem to take)? (Molitor, Ballstaedt & Mandl, 1989)

may either be in text or in pictures (or in combinations). If users have to switch between a textual presentation of the instructions and the visual presentation of the user interface (display), this also implies a switch from one medium (text) to another (visual).

It is difficult to bring the dual coding theory in accordance with the idea that we have to process both textual and pictorial information from both the spatial-graphical user interface and from the linear-procedural user instructions. Following the dual coding theory, we may perceive textual information on the user interface (e.g., display information) and visual information from the user interface, and store these separately. Then we switch to the user instructions, which also have both textually and visually presented information, and then we would separate these two and further process them separately to unite them again at some point to execute the series of acts on the user interface.

Another difficulty for the dual coding theory is the wide variety of types of pictures that exists. A picture may be a very realistic representation of reality in full color and full perspective. But it may also be a somewhat abstract black and white line drawing, or an icon or even a symbol or a diagram. This would call for a multi-code theory rather than a dual code theory, but it is also hard to see when exactly a representation is a picture at all (as for instance with schematic diagrams and flowcharts). Moreover, we also receive instruction via moving pictures: animations. Modern user interfaces consist of elements with many types of icons and symbols and in the accompanying help system there may be both realistic pictures and abstract pictures, perhaps presented in a flowchart model, and the instructions are sometimes presented in animations.

It is easier to believe that we have one single, amodal system in our minds to process the information presented to us via either the user interface (or deduced from our interactions with the user interface) and via the user instructions, whether this information is presented in pictures, text, diagrams and/or animations.

(A) Congruency and incongruency between encoding and reproduction media

An interesting study concerning the switch from text to visual or vice versa and its computational implications for the brain has been done by Seel & Strittmatter (1984). They presented subjects with either a written path description or a map with a drawn route, and asked them to reproduce the information in the other medium (subjects who received a map had to write a path description and subjects who received a path description had to draw a map). Two control groups had to reproduce the information in the same medium: one group text into text and the other group map into a map. Seel & Strittmatter found that congruency between the encoding and reproduction media leads to better performance[54]. If we can translate these conclusions to the human-machine interaction, with users interacting with the machine by using either text or using visuals as input to manipulate a predominantly visual (graphical, spatial) user interface, this implies that it would be easier for users to receive visual instructions, because the presentation of the instructions in the help would be congruent with the information presented by the graphical user interface: users then do not have to convert the information from one medium to another during the switch from instructions to machine. Moreover, visual user instructions present direct recognition (which text never can do). If the information is presented in text, on the other hand, the user always has to transform when moving from the instructions to the (spatial graphical) user interface. This means that the information in the instructions would be informationally equivalent (having the same content), but might be different computationally (the quantity of work for the brain).

---

[54] In their fine introduction to *Knowledge acquisition in text and pictures* (1989), Molitor, Ballstaedt & Mandl comment that the findings by Seel & Strittmatter are not as consistent or as interpretable as the authors suggest, but they do accept this conclusion, especially since it confirms the expectations.

This would make an easy hypothesis. But there is one more, vital aspect in the human-machine interaction involving additional help information. As has been stated above, the user develops an abstraction of the procedures he has to perform, because of the resemblance of the different procedures (patterns). Users will indeed already have an a priori idea of the procedures to be performed on the basis of interactions with machines that they believe to be comparable. They may have an a priori (abstract) procedure model. If so, this influence might well have the opposite effect: it would be easier (involving less computational burden for instructions that are informationally equivalent[55]) to receive instructions in text (propositional form), because texts are already a form of abstraction, whereas pictures may be exact replicas of the buttons that the user has to press (especially if he has to press virtual buttons on a screen) and can never be abstractions.

(B) Congruency and incongruency in user instructions

This long digression into various aspects concerning the human-machine interaction including instructions in a way brings us to the central research theme: the question concerning the relative advantage of various presentation media (text, pictures or animation) in instructions for operating complex technological products. The research questions in the following chapters have been defined to compare the relative efficiency of text and pictures in user instructions, especially concerning specific types of information. However, the results may also shed some light on the question whether users of complex technological products develop a procedure model of a device in text or in pictures and whether this varies dependent on the presentation of the instructions in either text or pictures. In more specific terms, the question may be described as follows. What is easier for users operating a device: receiving the procedural instructions in a presentation that allows direct recognition (as representational[56] pictures do) or receiving this information in a presentation that is incongruent and does not allow direct recognition,

---

[55] Two representations are informationally equivalent if any information that can be handled by the one representation can be converted into information in the other representation and vice versa. Two representations are computationally equivalent if they are equally useful and efficient. A representation is a format for recording, storing and presenting information together with a set of operators for modifying the information (Tabachneck-Schijf & Simon, 1996, pp. 28-29; the terms were introduced by Larkin & Simon in 1987).

but may be more in line with the abstract procedure model that users develop when using the various functions of the same device?

Representational pictures allow for direct recognition, but the inference as to what to do may be more difficult compared to text; we have to translate the pictures that present operational information into the instruction: 'What do I have to do?' If the instructions are offered in a propositional way, there can be no direct recognition of the elements, so the user has to make a small translation on this very elementary level; this implies a little more cognitive load. But the text is already a somewhat abstract pattern and further abstraction (the development of a procedure model) seems easier (less cognitive load). The abstraction leading to the procedure model:

[program] [*select sound variation*] [*select number*] [program],

is rather easy to represent in text as is shown externally in this sentence. It is hard to imagine such an abstraction as a procedure model in pictures, icons or symbols, because these would have to be abstract pictures, for instance, for 'select sound variation'. Such abstract pictures in the mind are hard to imagine. Therefore, it is unlikely that we make such a kind of translation. In case of abstraction to a more abstract procedure model, we must end up in one representation: text.

---

[56] Pictures that are identical or close to the original. In this study this means, for instance, that a picture of a button has approximately the same size, color and relative position to other elements of a display.

CHAPTER 6

Research into presentation media for product interaction

[This chapter is based on Westendorp, P.H. (1994), Design concepts of user manuals. In: Steehouder, M. et al., Quality of technical documentation. Rodopi Amsterdam/Atlanta, pp. 39-48.]

6.1. Comparisons of efficiency of various media

6.1.1. Variation

If there is one thing that can be concluded from the overview of the presentation media for product interaction (chapter 2), it is that designers, writers and artists have produced an overwhelming variation of types of instructions. They have presented a wide variation in concepts, layouts, instructive elements, texts and pictures to explain and instruct. They have made choices about whether to use texts or pictures or combinations and decided whether to use instructional elements for every detail instruction and how to combine all of these in concept and layout. The variation indicates that there seems to be little agreement about what is in general the best way to present instructions, or whether every specific situation has its own best way of presenting the instruction. Yet, there seems to be ample evidence that such factors do influence the efficiency of these aspects of user instructions (Van der Waarde, 1993, p. 116). It is especially the choice between texts and/or pictures in specific instructions that is of interest here, including the possibility to animate these.

6.1.2. Research

This question has been the basis of quite some research over the past decades. Overviews by Levie & Lentz (1982), Goldsmith (1984), Schriver (1989 and 1997), Wright (1988), Sims-Knight (1992), Van der Waarde (1993) and others present a wide variety of studies concerning the efficiency of the two media. Most of these studies compare the efficiency of the two media for training situations, textbooks and schoolbooks; the interest for interaction with a specific product is more recent. The conclusions of most studies pretend to have a practical value: it is advised to use pictures or texts in specific situations. In most of these studies pictures are considered as a supplement to text; a conclusion may, for instance, be that putting pictures in a text leads to more interest from students. It is remarkable how little the possible theoretical relevance is stressed; only very few studies seem to have tackled the way people learn (abstracting, generalizing), remember, decide, interpret or interact with products as the main goal and use of the various media as a means to find out more about these phenomena. On the other hand, schemata theories (John Locke; Schank & Abelson, 1977), mental model theories (Johnson-Laird, 1983; De Kleer & Brown, 1983; Kieras & Bovair, 1984), congruency or incongruency between coding and decoding (Seel & Strittmatter, 1989) and other theoretical issues have been used for the study of the relative efficiency of text versus pictures, but the results hardly ever seem to have been applied.

6.1.3. Criticism

Research concerning the efficiency of text versus pictures has met increasing criticism. Schumacher & Waller (1985) already found that in most tests both content and form of the material had been chosen subjectively. Stone noted back in 1980 that the materials used in the tests, where the effectiveness of text was contrasted with the effectiveness of pictures, were rarely discussed in terms of their relevant characteristics. This criticism was continued by Stone & Glock (1981) and Bieger & Glock (1984/5). Clark (1983) reconsidered the research on learning from different media and concluded that not much could be concluded on the basis of the studies he had seen. Winn (1989) discussed the methodological difficulties in testing the efficiency of graphics alone and concludes from a survey that the research on instructional graphics for particular tasks and student characteristics cannot be done unless the designer knows how graphics convey their meaning, that is their 'syntax'. 'Attempts to describe such a syntax have been few and insufficiently comprehensive' (p. 126). Wright (1988), Molitor, Ballstaedt & Mandl (1989), Sims-Knight (1992) have all criticized the unscientific setup of most of the studies. Their comments will be discussed here at somewhat greater length. In general, the criticism focuses on the fact that text and pictures have been compared without any description in generalizable terms. Therefore, the conclusions can only concern the specific situation (texts, pictures, product) and cannot be generalized. This means that such studies have little theoretical relevance, and are also most unlikely to have practical relevance: the conclusions are of little use to somebody who is to make user instructions for another product, target group or situation.

Molitor, Ballstaedt & Mandl (1989) conclude 'that the effect-oriented research is generally unsatisfactory, revealing a few serious methodological and theoretical deficiencies'. They have discussed the reasons at length, but here a summary will do. 'Indeed, various types of texts and pictures have been examined, but both forms of media have often been described either poorly or not at all with respect to their important cognitive aspects (for example, their complexity, organization, or sequencing). Worse still, systematic control was lacking with regard to the contents and formal relationships between text and picture.' [...] 'The experiments are usually formulated into ad hoc questions, and are not grounded in a cognitive processing theory. The results of applied research are therefore only helpful in practice to a limited degree, although much effort has been invested. The media designer is confronted with a patchwork of findings, which hardly provide recommendations for the respective design problem. The conclusion shows again that the rapid search for technological solutions in psychology is mostly condemned to failure' (p. 27).

Wright (1988) notes the wide variety of functions and media[57] for instructive documentation. Therefore, Wright continues, 'a broadly-based theory of effective communication would undoubtedly be an asset in knowing how to make decisions about documentation.' But,

---

[57] Wright uses the term 'medium' to refer to either paper or on-line; she uses the term 'modality' to refer to visual or auditory information and the term 'representation' to refer to texts, pictures and diagrams. Many other psychologists do as well, but often definitions of this term include either paper or on-line as well as pictures, texts or diagrams (see, for instance, Tabachneck-Schijf & Simon (1996, p. 28): 'A representation is a format for recording, storing and presenting information together with a set of operators for modifying the information'). In the present study the term 'media' is used in the sense of Molitor, Ballstaedt & Mandl (1989, p. 3): 'Texts and pictures can be regarded as media because they visualize parts of reality which are not immediately present or accessible in one's direct experience'. This allows the distinction between 'representational pictures' and 'non-representational pictures' (that cannot be perceived under normal visual conditions, like hierarchies, flow charts and organizational schemata). Both uses of the term 'media' can be found in literature; nowadays the term 'media' seems to be used more often than the term 'representation' to refer to text or pictures, especially in research concerning text, pictures, sound and animation (see, for instance, Feiner & McKeown and other articles in Maybury, 1993). Of course, this has to do with the use of the term multimedia to refer to the combination of text, graphics, sound and video/animation.

she immediately continues, 'Unfortunately such a theory does not yet exist, nor are we even close to producing it' (p. 629-630). Wright notices severe difficulties where the efficiency of presentation of instructions on paper or on-line is concerned. 'Empirical comparisons between on-line and hard copy documentation are inevitably fraught with confounding factors. Many aspects of the presentation change as the medium changes. When the design space is so large, how would one know if either was the best of its kind? And what useful purpose can be served by comparing a good example of X with a poor example of Y?' (p. 637-638). This refers to the comparison of paper versus on-line presentation, but the same is true for the comparison of instructions in pictures versus instructions in text: What is the use of comparing a bad text with good pictures or vice versa? As long as we do not have a generalizable definition of quality for texts and pictures, there is little use in making comparisons.

Wright also notes that generalization of conclusions concerning the design of a user manual are difficult, because 'the total design is a cluster of interacting elements. So guidelines from research about the format of particular elements may be overridden by other aspects of communication' (p. 640). To translate again towards the issue we are considering here: even if a valid comparison could be made between efficiency of pictures, it would be difficult to apply this in real world user manuals, because the text and pictures will be presented within a certain concept (as discussed above).

Sims-Knight (1992) wonders why so many visual representations are ineffective or worse. She too is convinced that 'the process of conceptualizing and creating visual representations rarely is grounded in knowledge that goes beyond the intuitions of the designer' (p. 325). Sims-Knight offers several interesting explanations. The first explanation is that in many studies a comparison is made between an original version using one medium (e.g., text) and a new version using the other medium (e.g., pictures). Sims-Knight (referring to Hartley) rightly remarks that 'it is easy to assume that because you have done something, it must be an improvement over earlier versions, whereas in actuality the change may not be for the better' (p. 329). One could add that it is always relatively easy to improve a text or picture that already exists and the improvement cannot always be contributed to the choice of the other medium; it might just as well have been possible to improve the version of the instructions using the same medium, but modify it[58]. If you change an existing text or picture you proceed on decisions already taken, and perhaps you would have made many other mistakes if you would have to start from scratch yourself.

Sims-Knight also sincerely criticizes the research into the relative efficiency of text and pictures in instructions. She argues that 'it is not possible to test a general hypothesis on the basis of specific examples in a specific context' (p. 347). This supports what Molitor, Ballstaedt & Mandl concluded in the text fragment cited above. Sims-Knight specifies with an example:

> 'In all studies of human behavior one may wish to test a general hypothesis, such as 'diagrams facilitate learning'. To test such a hypothesis, one must, however, implement a specific situation. One must use a particular diagram in a particular setting and one must use a particular test of learning. Even if the specific experiment is replicated several times in two or more laboratories, it may only be true within the narrow confines of its particular features. The hypothesis may not be supported when a different kind of diagram in a different learning setting with a different kind of learning is used' (p. 347).

Sims-Knight concludes that 'most of this research is atheoretical. It typically starts with a study that compares a text with a visual to a text without' [...]. 'Often the answer is complex - what is true for one kind of visual is not true for another, etcetera' (p. 351). In fact, Sims-Knight concludes that scientific research comparing the efficiency of texts versus pictures is of no use to designers ('Designers will often find themselves in situations in which scientific principles are not sufficient or applicable.'). She argues that user-based design, with 'quick and dirty' testing and redesign, is the only answer (p. 373).

---

58 This leads to a serious criticism of John Carroll's early work concerning the minimal manual, in which he simply improves existing material according to his minimal manual principles. His improvements, first of all, concerned more than the four minimal manual principles and secondly, even if only these principles would have been followed, they are simply improvements over an existing text, not the basic design principles of a totally new instructional text (see Carroll's examples in *The Nurnburg Funnel* and in his contribution in Doheny-Farina, 1991). Carroll's example of a minimal manual text can be strongly improved (Westendorp, 1995a and Westendorp, 1995b).

6.1.4. Generalizable terms

Overviewing this criticism, we may conclude that the problem is how to describe specific instructional situations in generalizable terms. The criticism also indicates that it will be very difficult or impossible to draw any generalizable conclusions concerning the relative efficiency of text, pictures or another medium in general. The best obtainable is generalizable conclusions for specific instructive actions in specific situations. As Bieger & Glock (1984/5) have indicated, describing specific instructive situations means that we have to try to describe the information content needed to perform that action in generalizable terms[59]. Of course, this limits possible conclusions to the level of the specific types of information in specific instructions. So the next question is: how to specify the information content for various instructional situations in generalizable terms?

6.2. Relative efficiency of text and pictures for specific types of information

6.2.1. Bieger & Glock taxonomy

Bieger & Glock (1984/5) presented a possible remedy to some of the methodological problems mentioned above. Their research elaborates on studies by Stone (1980), Stone & Glock (1981), Stone, Pine, Bieger & Glock (1981). They too immediately recognize the problem as formulated by Stone (1980): 'A major criticism of past research on pictures and texts has been that the materials used in that research were rarely described in terms of their relevant characteristics' (p.5). Bieger & Glock suggest as a possible remedy 'the development of a taxonomy of categories of information to classify the content of such picture-text materials in a way that would permit generalizability to other materials.' But they immediately continue with the remark that 'there has been little research done in the area of identifying the information content of either text or pictures' (p. 68).

Bieger & Glock based their taxonomy on the one designed by Mandler & Parker, expanded by Mandler & Johnson[60]. This taxonomy identified

---

[59] A comparison could be made with the ideas presented by Arne Naess on argumentation theory. Naess (1966) concluded that discussions about the truth or validity of an argument can only be drawn once the opponents in a discussion have specified their arguments on the same level.

[60] According to Bieger & Glock, 'The one notable attempt to identify the categories of information available in a stimulus [...] a taxonomy of information contained in pictures.'

four categories of information[61]:

1. Inventory information, which specifies what objects a picture contains.
2. Spatial location information, which specifies where objects are located.
3. Descriptive information, which specifies the figurative detail of the objects contained in the inventory.
4. Spatial composition information, which specifies the areas of filled or empty space and the density of filled space.

Bieger & Glock commented that this taxonomy referred only to the information available in pictures and did not include actions or reference to what could be inferred to be happening in the picture. Therefore, they added relevant categories from semantic case roles introduced by Fillmore, and predicate relationships introduced by Grimes. Fillmore identified several cases that linguistic entities can occupy (e.g., 'agentive' and 'instrumental') and Grimes described predicate relationships that were adopted in various analysis models (like 'covariance relationships' and 'response relationships'). Using the Mandler and Johnson taxonomy as a base and adding the relevant categories from semantic case roles and predicate relationships, Bieger and Glock developed their taxonomy of the information available in picture-text instructions. Their first step involved the identification of the kinds of information that people used when carrying out an assembly task. They did this by having students perform two assemblies while being videotaped and while thinking aloud. This comparison resulted in the addition, modification, or deletion of several pieces of information for each set of instructions. The second step involved re-analysis of the modified instructions using a discourse analysis system. They first produced a list of propositions that contained all of the information necessary for the assembly and then attempted to classify each proposition according to one of the categories described by Mandler & Johnson, or they defined new categories. This resulted in the following list of nine categories and definitions[62].

---

[61] In this description of the study of Bieger & Glock, many sentences of their text are copied literally without '' marks, because this would make the text needlessly complex.

[62] Bieger & Glock only gave examples in text; pictorial examples of these categories (slightly altered) can be found in Mijksenaar & Westendorp, 1999.

1. Inventory - information that specifies what objects or concepts are depicted. This information usually consists of the names of objects or concepts or the pictorial representation.
   Example in text: 'Connect *three large blocks* and *a small block* end to end[63]'.
2. Descriptive - information that specifies the figurative detail of the objects or concepts depicted. This information describes what the object looks like.
3. Operational - information that directs an implied agent to engage in a specific action.
   Example in text: '*Connect* three large blocks and a small block'.
4. Spatial - information that specifies the location, orientation, or composition of an object.
   4a. Location - describes the position of an object in space in relation to another object or fixed point of reference.
   Example in text: 'The large block *beneath* the small block'.
   4b. Orientation - describes the orientation in space of an object.
   Example in text: 'The end of the block is *pointing up*'.
   4c. Composition - specifies the area of filled or empty space and the density of filled space.
5. Contextual - information that provides the theme or organization for other information that may precede or follow it. This information typically specifies the general outcome of following certain procedures. Often it is a depiction of the finished product, but this information may occur at several levels, so it may be the outcome of an action at a lower level (like combining elements to construct a detail that consist of several elements).
6. Covariant - information that specifies a relationship between two or more pieces of information that vary together, such as cause and effect, problem and solution, goal and result. Bieger & Glock do not mention it, but probably covariant information is usually the feedback information.
   Example in text: 'Connect the rod and the clip *so that* the clip is in the middle of the rod'.
7. Temporal - information about the time course of states or events. This could consist of numbers (e.g., in pictures) or words ('*next*', '*then*', '*finally*')

[63] The italics indicate the specified information; in this case, the inventory information.

8. Qualifying - information that modifies other information by specifying the manner, attributes, or limits of that information. Example in text: 'Arrange the columns so that they are *about* two blocks apart'.
9. Emphatic - information that directs attention to other information. In pictures this information is for instance presented in bold lines, arrows or colors; in text emphatic information can be expressed with underlining or phrases like '*be sure that*' or '*notice*'.

Bieger & Glock then trained two new raters in the use of the categories and asked them to independently assign each proposition to one of the categories. This resulted in a high degree of agreement among raters. Further analysis of the categories of information and of the test results revealed that four categories of information were present at almost every step of the two assembly sequences: inventory, operational, spatial and contextual. The other categories of information occurred with a much lower frequency. Bieger & Glock hypothesized that the four frequent categories contained the essential information. They designed a test with all possible combinations of these information types (but for practical reasons excluded inventory information), presented either textually or pictorially. Subjects had to construct a model loading cart or fold a piece of paper into a certain figure, using one of the 36 total combinations of information types that resulted. The results of this test revealed that indeed operational, spatial and contextual information were the important and perhaps essential information for the completion of the assembly tasks. Subjects receiving the complete instructions completed the assemblies in significantly less time and with significantly fewer errors than those using incomplete instructions (i.e., missing one of the important categories). In this test the inventory information (e.g., parts' names and identities) was presented and trained separately before the test. Bieger & Glock concluded that the taxonomy might be a functional classification for describing the information content of procedural instructions. They also concluded that the three categories of information conveyed are very important, and that it is most beneficial to limit the information to these categories. It should immediately be mentioned that the fourth category ('inventory') also seems to be vital, as all subjects had a special training with this information previous to the test. Bieger & Glock indeed concluded in their follow-up article (1986, p. 181) that this type of information is also vital.

In a second study, Bieger & Glock (1986) applied their taxonomy of information types to compare the efficiency and effectiveness of text or pictures for the specific types of information that they had found to be vital for instructions. They varied the presentation of the information for the same tasks (constructing a model loading cart and a felt task[64]), using either pictures or text for the operational, contextual and spatial information respectively. Subjects performed tests, and their speed and accuracy were measured. The results indicated that textual presentation of spatial information produced fewer errors, while pictorial presentation of spatial information reduced performance times dramatically. It was further found that pictorial presentation of contextual information substantially reduced the number of errors (but there was no significant different as far as speed was concerned). Bieger & Glock found no differences between pictorial and textual presentation of operational information.

6.2.2. Criticism

'Further research must proceed on two fronts: a reliable taxonomy of verbal and spatial tasks (or a verbal-spatial continuum) must be developed to validate and generalize the principles of compatibility', Wickens, Sandry and Vidulich wrote in 1983 (p. 246). It seems that Bieger & Glock have come up with exactly such a taxonomy. Bieger & Glock have based their taxonomy firmly on quite generally accepted theoretical criteria and classifications (Fillmore; Grimes; Mandler & Johnson). They have also done considerable work to make sure the categories they finally produced are indeed the only types of information that are found or should be found in specific user instructions for assembly tasks. Nevertheless, criticism remains possible. Why, for instance, did Bieger & Glock not refer to Boohers' famous study (Booher, 1975)? Booher was one of the first to compare the relative efficiency and effectiveness of pictures and text in instructions applying some general categories, and he was certainly one of the most widely cited. He distinguished context, focus and action-step information types. Wickens, Sandry and Vidulich (1983) did a comparable study, and Bieger & Glock did not refer to this study either. That was perhaps because Wickens, Sandry & Vidulich only distinguished two types of information: verbal and spatial.

A more recent alternative is the classification of information in

[64] The felt task consisted of folding a napkin.

software manuals as just 'procedural' and 'declarative', as made by Ummelen (1997). This division of types of information is based on the division of memory into declarative and procedural memory. Unfortunately, this simple taxonomy is not useful for a contrastive analysis of the efficiency of text and pictures in user instructions. Too many parts of the information presented in user manuals could be categorized in each of these two categories, so this classification is too vague to make generalizable conclusions. Warnings can be classified as descriptive or procedural. Spatial information can be a description (where are the elements located?), but it may also be part of a procedure (where is the element that has to be pressed in this step of the procedure?). Covariant information can be vital in a procedure, but it may just as well be part of a description of the way the device reacts in general. For comparing the relative efficiency of text and pictures in user instructions for specific products, we need a more detailed specification of types of information.

Other criticism could refer to the design of the test (e.g., the 'think-aloud' procedure), the training of the raters (the training itself may easily influence the later rating of the information types) or the leaving out and training separately of inventory information. Another point of criticism refers especially to the textual presentation of operational information, but is probably also valid for the pictorial presentation: temporal information is often implied, simply because of the order of the sentences (or pictures).

In addition, a closer look at the pictures in the appendix of the Bieger & Glock 1984/5 article may lead to some question marks (figure 6.1). The pictures (to a lesser degree the texts) seem somewhat arbitrarily chosen. Other pictures might have produced different results. For the construction of the model loading cart, Bieger & Glock present a drawing of the four necessary blocks in arbitrary positions (the spatial information adds where the blocks should be). But can an arbitrary picture of four items that have to be combined be considered as operational information, information that has to inform the user what to do? Operational information would in this case - as in many others - be conveyed with instructive elements, in casu with arrows indicating how to combine the elements. Just by themselves, the picture of the four building blocks holds no operational information at all. No serious designer of user instructions would present only this operational information.

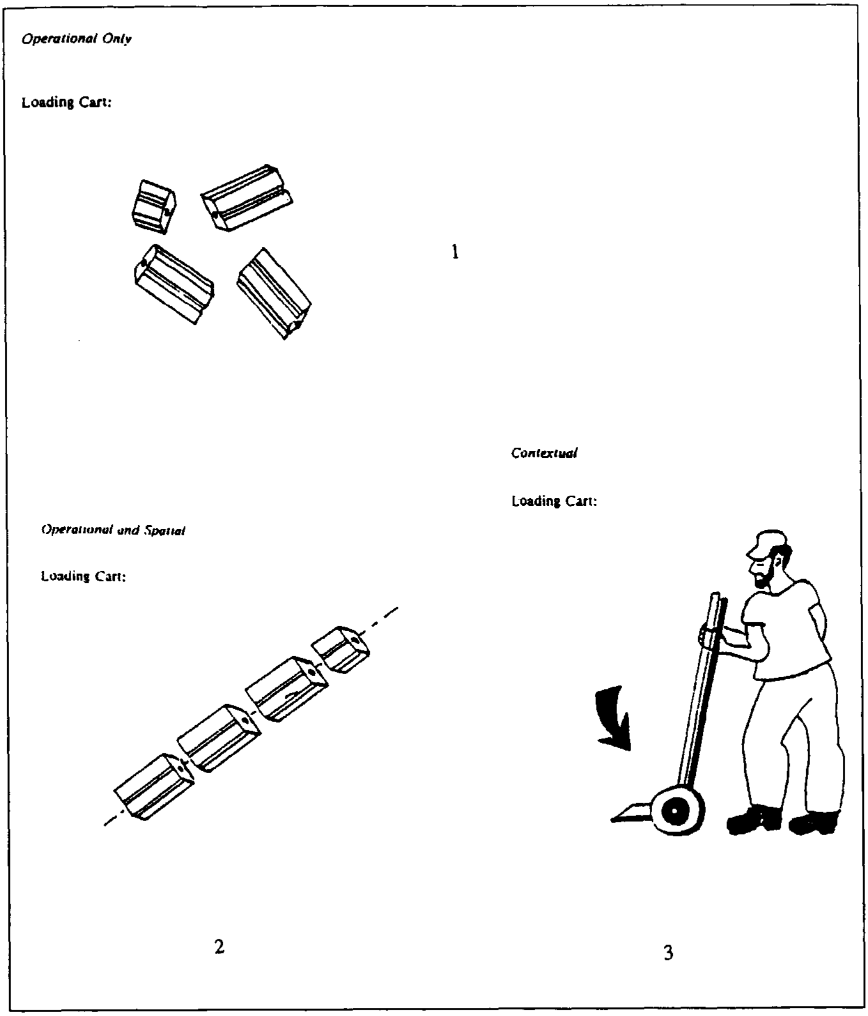


Figure 6.1. Bieger & Glock (1984/5 appendix). Some illustrations used in the instructions. Operational information only (1); operational and spatial information combined (2); contextual information (3). Note that these three illustrations do not indicate a procedure, but just types of information.

A problem for the application of the Bieger & Glock taxonomy of information types is the decision at which level the material should be analyzed. Bieger & Glock analyzed texts on word level and sentence level for all categories except contextual information, and pictures on the level of each picture as a whole. For the following tests it was decided to categorize according to paragraph and complete picture level, because it was often impossible to decide whether a sentence could be classified into one category or the other without deciding about the previous and next ones.

In spite of all criticism, the taxonomy that Bieger & Glock presented seems the best classification that we have at this moment. Anoshkey &

Catrambone (1992) compared their taxonomy with those developed by Booher (1975), Wickens, Sandry & Vidulich (1983) and Feiner & McKeown (1990). This comparison shows that the Bieger & Glock taxonomy is the most detailed one and that the Feiner & MacKeown taxonomy can be mapped on the Bieger & Glock taxonomy. Anoshkey & Catrambone conclude that the strong correspondence between the two most detailed taxonomies supports their use as a classification scheme. Moreover, the Bieger & Glock taxonomy has been applied by other researchers (Williams, 1993, Andre & Rist, 1993, Maybury, 1993) and in a training book for user manual writers (Kösler, 1990). Their taxonomy has been used and mentioned in several other studies (for instance, Schriver, 1997; Maes, Ummelen, Hoeken, 1996). Because of this circumstantial evidence of the usefulness of the taxonomy of types of information in instructions, it seems worthwhile to try to find out whether this taxonomy is indeed correct and applicable to other situations.

6.2.3. Application of the taxonomy of information types for products with 2D interfaces

It would be useful if we could indeed apply the Bieger & Glock taxonomy of information types to other types of products, especially for products with a two-dimensional interface including VCRs, microwave ovens, software and other modern electronic household and office products. It is especially these products that confront the users with many problems; to get these products to function, users have to rely heavily on the user instructions. These products probably pose many more problems to the users than 3-D-assembly products do. The user instructions for these products also consist of combinations of texts and pictures (and instructional elements), although there is probably relatively more text and there are relatively fewer pictures than in user manuals for the assembly of three-dimensional products.

In summary, modern electronic household and office products have rapidly become very versatile and, as a result, very difficult to use. Yet, at the same time these products often have displays and screens, so that a wide variety of feedback and on-line help can be offered. Compared to printed manuals, on-line help offers new possibilities. On-line help can be interactive and offer hyperlinks, influencing the structure and accessibility of the information (a test is discussed in chapter 10). On-line help also allows help in other media than text and still pictures: animation, video and sound can also be applied. Of these three, especially animation might be helpful in the product's user instructions. Video has the same advantages and especially disadvantages as photography compared with drawings in user instructions on paper: they show far too many details, including all the irrelevant ones. Sound can be very useful for feedback, especially the sound that the apparatus can, should or should not make. It is questionable whether the human voice will be very useful for user instructions for specific products. The disadvantages are well known: a recorded human voice is very much a linear presentation, and therefore in practice often too slow, especially if the user wants to repeat the instruction. Of course, in some situations a recorded human voice is ideal for instructions, e.g., for pilots or surgeons in action or for en-route information systems in cars. Animations can be compared with drawings; in fact the only difference is movement. Therefore, it could be useful to try and find out whether the taxonomy of information types in instruction can be applied to select for each specific instructive situation, not only between text and pictures, but also between text, still pictures and animations. Some contrastive tests, based on specific hypotheses derived from the more general ideas presented in this chapter, will be presented and discussed in chapters 7, 8 and 9.

CHAPTER 7

Text, pictures or flowcharts for specific types of information in instructions for two-dimensional products

[This chapter is based on: Westendorp, P.H. (1999a), *Text, pictures or flowcharts for specific types of information*. In: Zwaga, H., Boersema, Th. & Hoonhout, H. Visual information for everyday use. Taylor & Francis, London, pp. 83-89.]

7.1. Introduction

As has been argued in the previous chapter, a classification of the types of information that may occur in user instructions is a prerequisite to examine the relative efficiency of text or pictures in user instructions and to draw generalizable conclusions. Bieger & Glock designed a taxonomy of nine different types of information in user instructions concerning assembly tasks and this taxonomy seems to be the best available overview of types of information in user instructions, in spite of some criticism. This taxonomy allows classifying information presented in user instructions and performing contrastive tests, varying the medium in which a specific type of information is presented. On the basis of such tests, conclusions can possibly be drawn concerning the relative efficiency of a presentation of that type of information in either text or pictures.

2-D product instructions: more often and more important

The Bieger & Glock taxonomy was designed for assembly tasks only. It was tested with the instructions for assembling a model loading cart and for doing a felt-task. The taxonomy clearly refers to tasks to be performed with three-dimensional products. At the present time, however, a major part of the instructions concerns two-dimensional products. Moreover, the latter instructions seem to bother people most: the instructions that accompany the VCR, the word processor, the car-navigation system, the telephone and other modern technological products. Therefore, it may be useful to test this taxonomy for products with a two-dimensional user interface. Such products may be either software (completely two-dimensional) or three-dimensional products with a user interface that only has two-dimensional aspects for the user, such as modern VCRs, route-information systems, cameras, telephones, fax machines, copiers and control systems (telephone switches, power plants). The criterion for deciding whether we deal with two-dimensional interfaces is that the user has to deal with a more or less two-dimensional display, usually consisting of buttons, displays and screens. For the instructions there is not much difference with completely flat user interfaces, such as the ones that are used to present software. The user mainly has to press or touch real or virtual buttons and notice the reaction of the device.

2-D pictures: less arbitrary

There is another reason to try to apply the Bieger & Glock taxonomy of information types to instructions for two-dimensional products. Pictures of two-dimensional products or details can be less arbitrarily chosen than pictures for the instruction for use of products involving three-dimensional tasks. In three-dimensional pictures many choices have to be made that may have a severe impact on the efficiency of the picture and we are far from any standardization in the visual elements of user instructions (Dirken, 1999, p. 270). The illustrator, for instance, has to decide whether to make a perspective drawing (more realistic) or an axonometric drawing (more simple, measurable), to make a ghost view, a cutaway or an exploded view, choose a point-of-view and orientation out of an unlimited number of possibilities, to decide what details of the whole to present, etc. Bieger & Glock did not motivate any of these choices concerning the pictures that they applied in their tests. For the kind of equipment for which only the front view with keys, controls and a display or a screen is relevant, simple two-dimensional pictures are the obvious choice. Discussion of the form will be restricted to style elements, such as the thickness of the lines, the size, the use of color and grey-value, and the degree of abstractness of the pictures (very realistic or more stylized). Still quite a few variables remain that limit generalizability of conclusions in contrastive tests, but far less than in a situation with three-dimensional pictures (Van Mourik & Westendorp, 2000).

7.2. Applying the taxonomy to instructions for 2-D products

To examine whether the Bieger & Glock taxonomy could partially be applied to instructions for two-dimensional products, a test was done. Four groups of three students first studied the relevant Bieger & Glock articles (1984/5 and 1986) and received an explanation of the taxonomy. Also, some training exercises were done with user instructions of three-dimensional products. Then these four groups of students independently analyzed the instructions of a Lacis office telephone with many functions, applying the Bieger & Glock taxonomy.

Level of analysis

A major problem for application of the Bieger & Glock taxonomy is the decision at which level the material is to be analyzed (MacDonald-Ross, 1989, p. 148). Bieger & Glock noted this problem for contextual information only and decided to analyze contextual information on a paragraph level and all other information on word level (see the verbal examples in chapter 6 and the visual examples in Mijksenaar & Westendorp, 1999). In our test, though, analysis on sentence level produced highly different results from analysis on paragraph level. Therefore, we found analysis on two different levels for different types of information unsatisfactory. Moreover, the same type of problem occurred when trying to apply the taxonomy to the pictures: often it was found possible to attribute various types of information to one and the same picture. This is also true for the pictures that Bieger & Glock had used to represent what they considered to be a picture for one type of information (see Bieger & Glock, 1984/5, appendix).

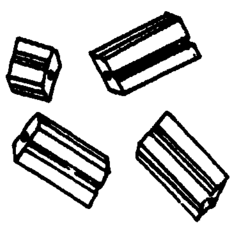


Figure 7.1. (left) Operational information in picture. From: Bieger & Glock 1984/5, appendix.

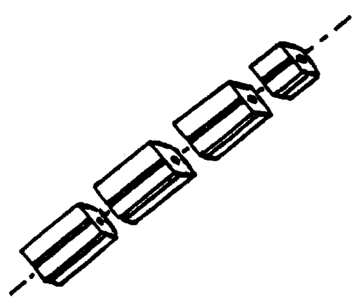


Figure 7.2. (right) Operational and Spatial information in picture. From: Bieger & Glock 1984/5, appendix.

For instance, in figure 7.1, according to Bieger & Glock only operational information is presented. However, it can be maintained that this picture also presents 'inventory' information: it gives an overview of the elements needed to assemble (this part of) the model loading cart. On the other hand - and perhaps more serious - one could question whether this picture does present operational information. Does this picture indeed 'direct an implied agent to engage in a specified action'? Is it the pictorial counterpart of the verbal example of operational information given by Bieger & Glock (p.70): 'Connect three large blocks and a small block'? This picture does not indicate that the user has to 'connect' something; it merely presents the four blocks that have to be connected. This seems to be an argument to only call this inventory information.

Likewise, in figure 7.2 Bieger & Glock present a combination of operational and spatial information. However, a close look indicates that in this picture too, other types of information can be discerned, apart from operational and spatial. We can discern inventory information, but also contextual information - because we do see what the final product of this phase is supposed to look like - and temporal information, because we do see in what order we have to connect the blocks.

In user instructions, normally the operational information is presented visually in the form of arrows, fingers, hands or dotted lines indicating what to do. Such instructive elements (chapter 2) often also include other types of information, for instance, the way in which elements have to be combined: spatial information (figures 7.3 and 7.4).

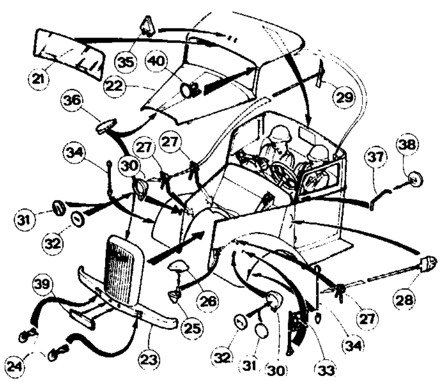


Figure 7.3. (left) Operational information and spatial information combined in the arrow and hands. Also note the covariant information (reaction: the paperclips falling out).

Figure 7.4. (right) Operational and spatial information in arrows. Also note inventory (numbered elements), contextual information (the exploded view is in a way a representation of what the finished product should look like), temporal information (if the numbers indicate an order of assembling).

Analyzing the existing user manual of the Lacis telephone, and applying the Bieger & Glock taxonomy on a paragraph (block) or complete picture level led to a very high degree of agreement among the raters (students). In fact, all groups came to the same classification and within the groups no important disagreements were reported. This result is in accordance with the high inter-rater agreement that Bieger & Glock reported: raters (also students) reached more than 97%

agreement about the classification according to taxonomy of the propositions in the instructions presented to them (1984/5, p. 69). Therefore we concluded that we could analyze the material using the Bieger & Glock taxonomy at a more abstract level.

Dominant: operational, spatial, covariant & contextual information

Further analysis of the user instructions of the Lacis telephone by the same groups of students showed that in this user manual too, four types of information were ubiquitous and the five others were reported to occur far less often, just like Bieger & Glock had found. In contrast to Bieger & Glock, however, we did not find inventory information ubiquitous, and we did find covariant information in many parts of the user manual. This result can easily be explained. In user instructions for two-dimensional products, it may be less useful to have an overview of the elements of the product (although this inventory information is presented quite often: the foldout page with an overview). In products that present feedback, there is likely to be mention of this feedback (covariant information) in the user instructions. As expected, just like Bieger & Glock, we found operational information abundant; a user manual without information telling exactly what to do is unthinkable. Also, not surprisingly, we found quite a bit of spatial information. Therefore, it was concluded that operational, contextual, covariant and spatial information were not only the most frequently occurring types of information in these instructions, but also the most important ones.

Now, it could be tested which medium (text or pictures) would be most efficient for each of these types of information in instructions. Time and number of participants limited the number of variations that could be made; somewhat arbitrarily operational, contextual and covariant information were selected for further testing. It was decided to include not only representational (realistic), but also logical pictures (flowcharts), because Wright & Reid (1973) found that flowcharts might be an efficient alternative to text in instructions for tasks that involved repetitive decision-making. The latter is characteristic for installation and use of two-dimensional (electronic) products. Especially in the case of user manuals for telecommunication products, flowcharts are applied abundantly (some examples are presented in Mijksenaar & Westendorp, 1999, pp. 100-101).

7.3. Test

On the basis of these findings and considerations seven versions of a user manual for the telephone (figure 10.3) were designed and compared in tests, in which text and representational or logical pictures for operational, contextual and covariant information were varied. Version 1 was a rewritten version of the existing user manual, with all operational, contextual and covariant information in text. The specified types of information in version 1 were presented in logical pictures (flowcharts), representational pictures or left out. It is obvious that operational information is vital for user instructions; therefore no version was designed leaving out this type of information. User manuals were designed with operational information in either pictures (version 4) or in a flowchart (version 5), so that efficiency of these medium choices for operational information can be compared to operational information in text (version 1). Furthermore, versions were designed with either the contextual information in flowcharts (version 2) or covariant information in flowcharts (version 6), as this seemed a more obvious choice to present these types of information than representational (realistic) pictures. One could question whether contextual and covariant information are indeed vital. Therefore versions were designed without these types of information (version 3 is without contextual information, version 7 is without covariant information). For an overview, see table 7.1.

Table 7.1. Seven variations of a user manual for a telephone, varying medium choice for contextual, operational and spatial information.

version	information type	medium
1	[existing user manual]	text
2	contextual information	flowchart
3	no contextual information	
4	operational information	pictures
5	operational information	flowchart
6	covariant information	flowchart
7	no covariant information	

Text was always in simple present, imperative style. Flowcharts were designed according to the ISO 5807 standard (ISO, 1985). Representational pictures were simple, realistic two-dimensional drawings of buttons, displays, etc. (the element of the telephone which was relevant at that moment of decision, usually presented in approximately actual size). The size of the user manuals varied from 30 to 35 pages. Details are presented in figures 7.5 - 7.9; the texts have been translated into English; the design, layout and typography are identical to the design, layout and typography of the existing Dutch versions that subjects have used. These variations of the user manual for the Lacis telephone were compared in contrastive tests.

Figure 7.5.
Detail of version 1: the existing user manual (text) [translated].

PROGRAMMING DSS-KEYS	
Press SHIFT and PROGRAM	the multi-line telephone is now in the programming-mode.
Press SHIFT and DSS-key	The content of the memory-place is empty; the display shows a dashed line.
Select a number or a code	The entered number or code is presented on the display. The number or code can be changed using the Delete-key.

Seventy subjects were asked to perform the same five tasks, and performance times were measured for each subject. For each version ten subjects performed the five tasks individually. All subjects were students at Delft University of Technology, male and female equally distributed over the seven versions; most were from the Faculty of Industrial Design Engineering, some from other Faculties, and most were 2nd or 3rd year students. The tasks had an order of increasing difficulty (table 7.2, p.149), so that subjects could get used to the telephone system and the user manuals. The easiest task was to program one memory key; the most complicated task was to set up a conference between two lines, excluding oneself. Each subject had half an hour for the five tasks; a pretest had indicated that this was enough. Performance times were recorded in seconds. For each subject the total time to complete all five tasks was computed; then, for each version the total time of the ten subjects was used to calculate the mean. t-Tests were conducted to analyze the scores for the seven versions. If a subject was unable to perform a task (or a certain step in the performance of a task), help would be provided after two minutes. The selected version of the user manual was provided for each subject, but subjects were not explicitly asked to read or use it. It was expected - and proved true - that all subjects would need the user manual to successfully complete all the tasks.

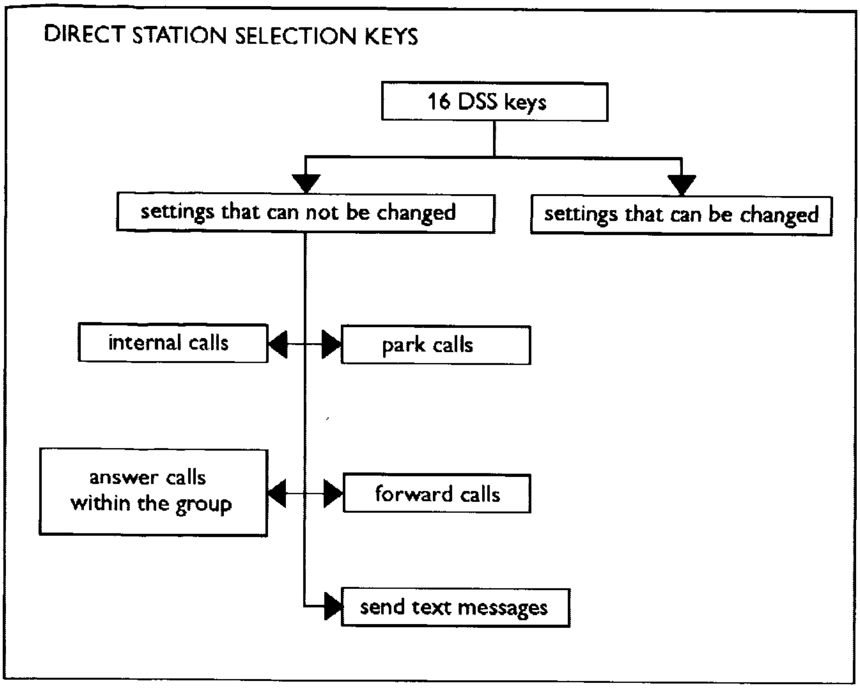


Figure 7.6.
Detail of version 2: contextual information in a flowchart [translated].

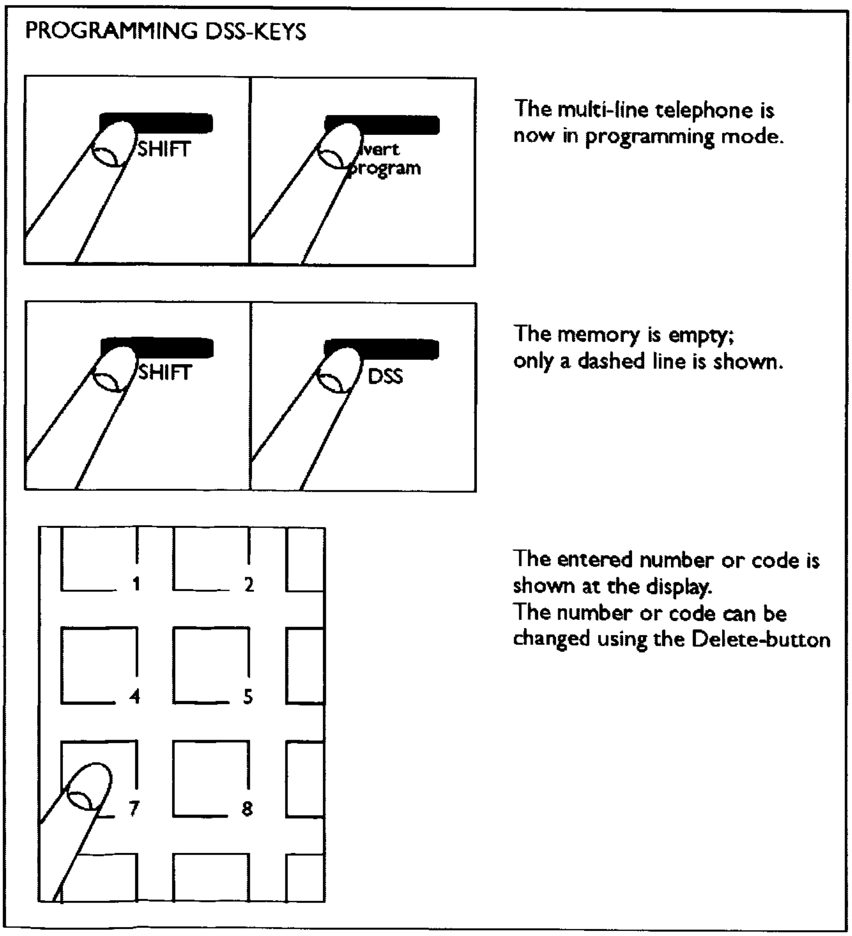


Figure 7.7.
Detail of version 4: operational information in representational pictures [translated].

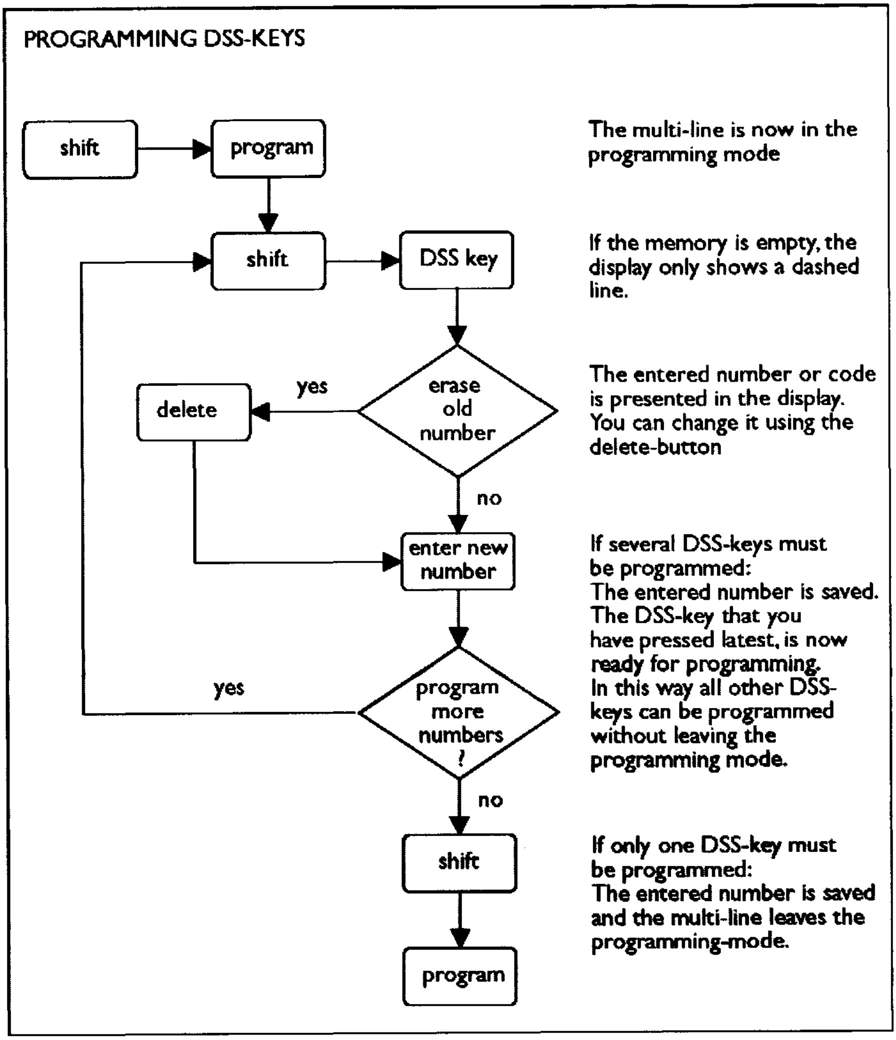


Figure 7.8.
Detail of version 5: operational information in flowchart [translated].

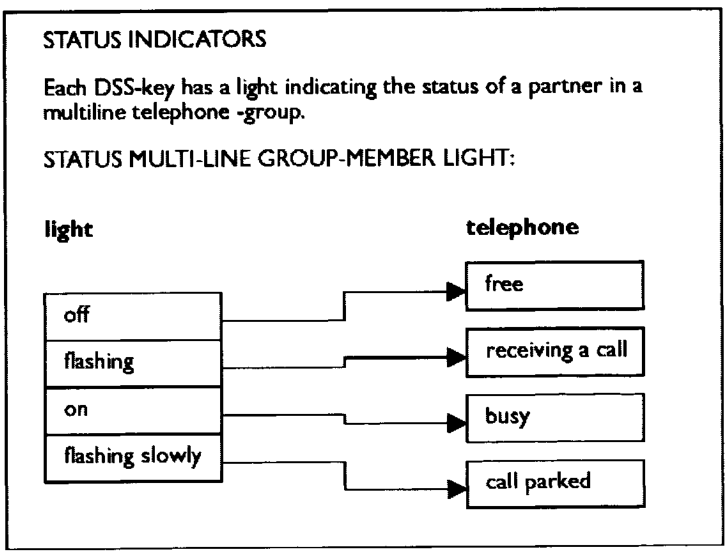


Figure 7.9.
Detail of version 6: covariant information in flowchart [translated].

Table 7.2 tasks that the subjects had to perform on the telephone [translated from Dutch]

1	Program DSS- key 8 with number 8970.
2A	Call this number, using the key you just programmed.
2B	The number is busy; program the phone so that it rings as soon as the number is free again and answer when the bell rings.
3A	Change the speech mode of the phone to loudspeaker mode (hands free).
3B	Call number 8970.
3C	Disconnect number 8970 without changing the speech- mode.
3D	Repeat the last chosen number and then disconnect.
4A	Call number 8970 again. You are now talking to the person at number 8970, but you want to consult somebody at number 8971.
4B	Call number 8971, but do not disconnect the line with number 8970.
4C	As soon as you have reached the person with number 8971, resume your connection with number 8970. Put down the phone.
5	Call number 8970 again. The person at the other end of the line now also wants to talk with the person with number 8971. Set up a telephone conference.

Performance times were considered as the only measure of efficiency, because making mistakes would cost time to correct and therefore be measured according to performance times. Moreover, it is hard to unequivocally decide what exactly has to be counted as a mistake and how seriously that has to be considered.

7.4. Results

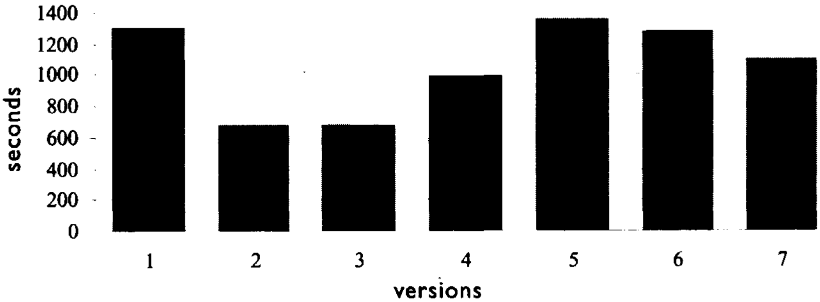
Detailed results (measured times per subject and per task, means, standard deviations and an analysis of variance) are presented in Westendorp et al. (1994); here only totals per version, means and standard deviations are presented (table 7.3 and figure 7.10).

Table 7.3. Mean total times that subjects needed to complete the five tasks successfully. N=10 per version, but in versions 3 and 4 only 9 subjects performed the tasks.

	version 1	version 2	version 3	version 4	version 5	version 6	version 7
Mean	1305	686	614	890	1362	1272	1104
Standard deviation	174	90	81	448	141	154	72

Figure 7.10. Mean times to complete the five tasks (in seconds) for the seven versions.

1 = Existing user manual (text);
2 = Contextual information in flowcharts;
3 = No contextual information;
4 = Operational information in pictures;
5 = Operational information in flowcharts;
6 = Covariant information in flowcharts;
7 = No covariant information.



For each version, ten subjects were asked to perform the five tasks, but in versions 3 and 4 only nine subjects actually performed the tasks, because of a misunderstanding with the test supervisors.

In this test, leaving out certain types of information (covariant and contextual) in instructions, and presenting specific types of information in different media (text, representational pictures, flowcharts) did lead to significant differences in performance times (one-way ANOVA: $F = 13.39$, $df = 6$, $p < 0.001$).
The versions with the contextual information in flowcharts (version 2) and without contextual information (version 3) proved to be significantly more efficient than the version with contextual information in text (version 1) ($t = 7.79$, $df = 18$, $p < 0.001$ and $t = 7.22$, $df = 17$, $p < 0.001$ respectively).

The experimenters supervising test version 2 (with contextual information in flowcharts) reported that this information had been used by only one subject for one task. Therefore, version 2 could - to a certain degree - also be considered as a version without contextual information. On the other hand, this means that not much can be concluded about the importance of having contextual information in either medium compared to the other. If we compare both versions 2 and 3 (without contextual information or with this type of information hardly used) with version 1 (which included the contextual information in text), it becomes all the more apparent that including contextual information seems unnecessary and even inefficient. So, in contrast to Bieger & Glock, we found contextual information superfluous. It could be that users working on products with two-dimensional user interfaces do not need to know a priori explanations of what will be the result of their actions. Discovering afterwards what exactly was intended was probably more efficient.

Presenting operational information in representational pictures (version 4) is significantly more efficient ($t = 3.03$, $df = 18$, $p < 0.008$) than presenting this type of information in text (version 1), but the variance between subjects is high (Levene's test for equality of variances: $F = 0.047$, $p < 0.830$). Some subjects did the tests quite quickly, but others performed much slower when using pictorial operational information. Bieger & Glock found little effect of the mode of operational information on performance.

Presenting operational information in a flowchart appeared to be counter-productive: version 5 was the least efficient of all. This is remarkable since all subjects were students of a university of technology and had experience with reading (and designing) flowcharts. Perhaps flowcharts are more useful for complex tasks with many decisions to be made. In this test, with its rather simple tasks, texts and representational pictures proved to be more efficient.

The existing user manual (version 1) contained quite a lot of covariant information. Comparing this version plus version 6, which had the covariant information in flowcharts, with version 7, which was without covariant information, showed that version 1 was significantly less efficient than version 7 ($t = 2.21$, $df = 18$, $p < 0.04$). Version 6 (covariant information in flowcharts) was also less efficient compared to version 7 (no covariant information), but the difference was not significant ($t = 1.14$, $df = 18$, $p < 0.176$).

It may be somewhat premature to conclude that covariant information is better left out for these kinds of tasks. In our tests the subjects needed the covariant information only for one of the five tasks. For this specific task the version without covariant information was more efficient than the versions with covariant information; subjects using version 7 (no covariant information) performed this task 38% faster than those using version 6 (covariant information in flowchart), and 32% faster than those using version 1 (covariant information in text). In our tests covariant information consisted usually of the feedback from the telephone (presented on the display or by a sound). It seemed that users did not feel the need to have any kind of confirmation in the user manual regarding the feedback of the machine. Either a sound or the information on the display was enough. This may be because the reactions of the telephones are generally known and understood. It may be different in situations where the user is left in uncertainty by the feedback of the machine. Another explanation might be that the subjects were experienced in using complex high-tech products. They may have more confidence in themselves than the average user.

7.5. Conclusions

This test indicates that representational pictures may be a relatively efficient medium for presenting operational information in instructions for products with a (predominantly) two-dimensional user interface. Presenting this type of information in either text or a flowchart leads to the least efficient instructions of all (versions 1 and 5 respectively).

The representational pictures in this test were very realistic, although black and white. They consisted of line drawings of the buttons that users had to press including the areas directly around them; they usually included a word (text) that indicated their function. Therefore 'pictures' here must be interpreted somewhat differently from the kind of pictures for three-dimensional products (like Bieger & Glock used) and conclusions can definitely not be generalized to pictures for operational information in instructions for three-dimensional products. In fact, one could argue that the word (text) was the most relevant part of the picture, since almost all the buttons were identical. Therefore, we could state that the comparison here is between direct recognition (an exact representation of the detail) and abstraction (the text). Presenting the operational part of the instruction in a medium that allows direct recognition, lead to a significantly faster operation of the product than presenting this information in a more abstract medium, like text or flowchart.

The instructional pictures could in this situation be chosen far less arbitrarily than in Bieger & Glock's test, but still some arbitrary choices had to be made. This did not so much concern the way in which the instruction was depicted, but especially what moment of the action was to be depicted. However, for these kinds of instructions this is again far less arbitrary than for instructions for three-dimensional products, because we simply had to present the button (area) that had to be pressed at every step of a procedure. Animations would reduce this problem further, because then all actions can be presented completely.

Flowcharts appeared to be an inefficient medium for the presentation of various types of information in instructions for two-dimensional products. Because the contextual information in version 2 (contextual information in flowchart) was hardly used, not much can be concluded concerning the efficiency of flowcharts for contextual information. The version with operational information in flowcharts (version 5) was significantly less efficient than the version with this type of information in representational pictures (version 4). Version 6, with covariant information in flowcharts, was also one of the least efficient user manuals. Version 5, with the operational information in flowcharts appeared to be the least efficient of all. This is important, because in the cases in which flowcharts are applied in user manuals, it is especially for operational information. See for instance figure 7.11.

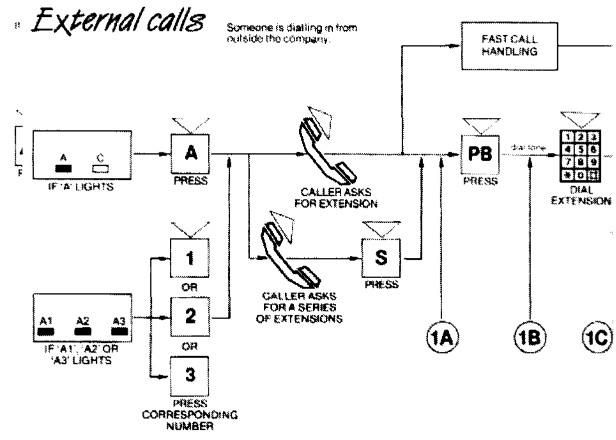


Figure 7.11. Detail from a manual for the operation of a Philips telephone.

Leaving out specific types of information seems to be quite efficient for user instructions for two-dimensional products. As has been argued above, version 2 (contextual information in flowcharts) could be considered as a version without contextual information (because this was - unintentionally - only used by one of the subjects for one of the tasks). The results show that both these versions (2 and 3) are very efficient (although it is somewhat more difficult to conclude this, because of some problems with the testing of these two versions). This indicates that it is more efficient not to include contextual information, in whatever medium, in instructions for two-dimensional products. Leaving out covariant information also seems more efficient than including it (at least in flowcharts). This may give support to the ideas Carroll has expressed about the minimal manual (1990). The findings reported here may also help to specify the guidelines for a minimal manual: what information exactly can be left out?

CHAPTER 8

Text, pictures or animation in user instructions

[This chapter is based on: Westendorp, P.H. (1996a), *Text, pictures or animation in instructions for use. A validation of different media for specific types of information.* In: Ensink, T. & Sauer, C. (eds.), Researching technical documents. Groningen series on language and communication. Groningen University Press, Groningen, pp. 181-197.]

8.1. Introduction

Modern technological products are very versatile and consequently very complex in use. In spite of good user interfaces, it is often difficult to instruct how to use all features efficiently and effectively. On the other hand, such products often have a display or a screen on which the instructions can be presented on-line.

Compared to printed manuals, on-line help offers more possibilities. It can be interactive (operate the machine directly from within the help system, see figure 3.4) and offer hyperlinks (to other information and demos). This, of course, may strongly influence the structure of the information. On-line help also allows help in other media than text and still pictures: video, sound and animation can be applied to replace text and pictures or as an addition.

In this chapter, the focus is on animation as an alternative to text and still pictures. Animations can be used to show how things work or should be operated, indicating the relevant details in their context. In contrast to still pictures, animations show every detail of a movement, which could be an advantage. Animations may be either additional to text and to still pictures or a replacement; the latter is subject of study in this chapter.

The possibility to apply animation in on-line help poses the question whether animation is indeed an efficient medium to communicate instructional information for the operation of modern technological products. As has been argued before, such a question cannot be answered in general. First of all, the efficiency can only be expressed relative to the efficiency of other media, such as text and still pictures. Moreover, generalizable conclusions concerning the relative efficiency of various media for instructing must be limited to specified types of information in these instructions (Bieger & Glock, 1984/5; see chapter 6). Contrastive studies applying text in one case; still pictures in one alternative and animations in another, to measure the efficiency with which subjects can operate the product, would never lead to generalizable conclusions if the type of information is not specified. The choice of which information would be presented in which medium in each design of the instructions would be too arbitrary, so that conclusions would be limited to the tested instructions. When animation as a medium is contrasted with other media, the Bieger & Glock taxonomy can be applied in the same way as to studying the relative efficiency of text and still pictures. The question then is restricted to: 'For what type of information in instructions is animation more and less efficient than text or pictures?'

Not all types of information will be considered; some are more important than others, just like Bieger & Glock have reported. What types of information seem the most interesting ones to select for a contrastive test, focusing on instructions for two-dimensional products, comparing text, pictures and animations? The previous test (chapter 7) indicated that contextual and covariant information may perhaps be better left out completely. No specific conclusions could be drawn concerning the relative efficiency of any of these media (text, flowcharts, realistic pictures) for contextual or covariant information. Therefore, these will not be included in the next contrastive test. It is obvious to include operational information again, since instructions without this type of information seem unthinkable and animations are often especially used to present operational information: show the actions to be performed.

Bieger & Glock found spatial information essential in instructions. In the previous chapter it was argued that spatial information seems less important for instructions for two-dimensional products. Spatial information consists of three subtypes: 'orientation', 'composition' and 'location'. For two-dimensional products 'orientation' and 'composition' indeed seem less relevant. Users of two-dimensional interfaces do not have to find out the orientation of elements on the product and neither do they have to compose elements. Presenting information about the location of elements, however, could be relevant, because users have to find where the buttons, icons, symbols and displays are located. Therefore in this study, the focus will be on the relative efficiency of text, pictures and animations for presenting operational information and on the relative efficiency of adding spatial (location) information in user instructions.

It is hypothesized that instructions for use are more efficient when operational information is presented in animation than in still pictures or text, and more efficient in still pictures than in text. It is further hypothesized that user instructions are more efficient when additional spatial information is presented (in pictures), following the results of Bieger & Glock (1986) and Westendorp (1999a, see chapter 7).

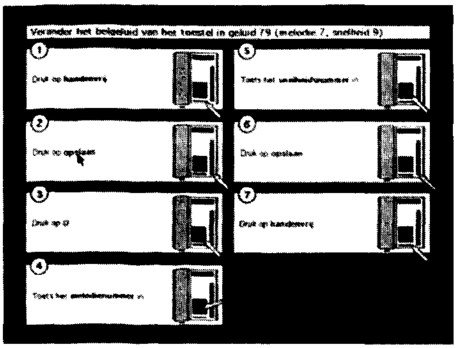
8.2. Method

To test these hypotheses, six variations of instructions were designed for the completion of a number of tasks to be performed on a telephone system. These included versions with operational information in either text, in still pictures or in animations; and each of these three versions was either combined with spatial information (presented in separate pictures) or without spatial information. This resulted in the following six versions:

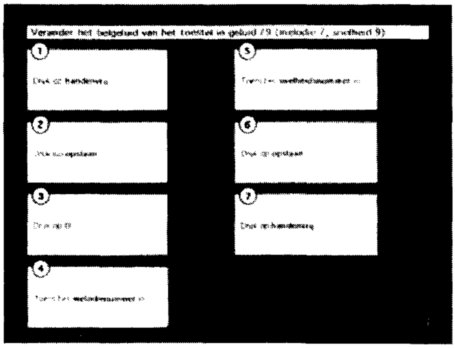
1. Operational information in text and spatial information (separate) in pictures;
2. Operational information in text and no spatial information;
3. Operational information in pictures and spatial information in (separate) pictures;
4. Operational information in pictures and no spatial information;
5. Operational information in animation and spatial information in (separate) pictures;
6. Operational information in animation and no spatial information.

Of each version one screen-dump picture is presented below (figures 8.1-8.6).

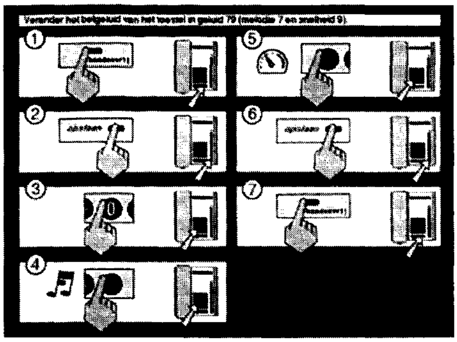
All instructions were designed and presented in MacroMedia Director 3.0, which enabled the presentations of animations on a computer screen. This software also enables the computer to both register the times that subjects needed to complete each task and calculate the total times. After each instructed task that was performed successfully (judged by the test supervisor), the subjects could click on the 'OK' button, whereupon the next task plus instructions were presented on the screen.



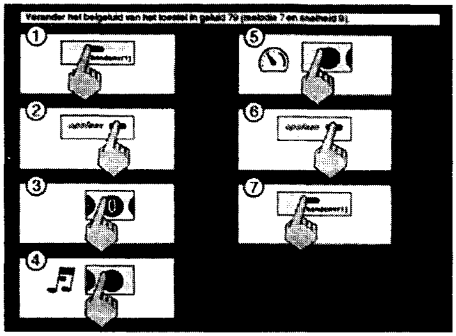
1



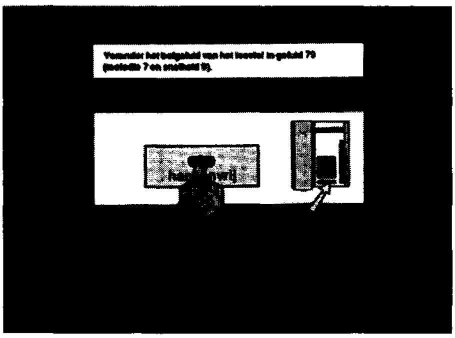
2



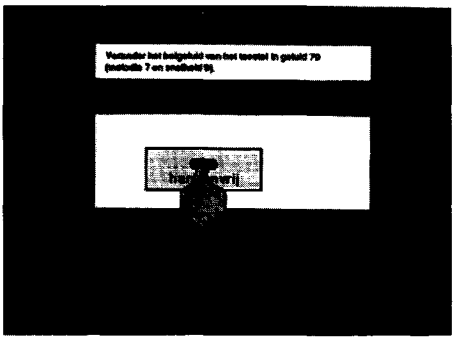
3



4



5



6

Figure 8.1. Operational information in text and spatial information (separate) in pictures. [Translation: Task: Change the bell-tone of this phone to tone 79 (melody 7, speed 9).

Verbal instructions:
1. Press hands free;
2. Press program;
3. Press 0;
4. Press the melody-number;
5. Press the speed-number;
6. Press program;
7. Press hands free.]

Figure 8.2. Operational information in text and no spatial information.

Figure 8.3. Operational information in pictures and spatial information in (separate) pictures.

Figure 8.4. Operational information in pictures and no spatial information.

Figure 8.5. Operational information in animation and spatial information in (separate) pictures. The hand moves towards the button; the button shows an effect as if it were pushed down; the hand moves away from the button and the button shows an effect as if it moves up again.

Figure 8.6. Operational information in animation and no spatial information. The hand moves towards the button; the button shows an effect as if it were pushed down; the hand moves away from the button and the button shows an effect as if it moves up again.

All elements were pictured as realistically as possible. Buttons and other elements of the telephone were depicted in the instructions in the colors that were identical or nearly identical to the originals, and the sizes were also almost identical or at least proportional to each other as they were on the real telephone. Typographical elements were also copied as closely as possible to the originals. Animations presented a move-in from the left of a picture with the key to be pressed (identical to the picture versions) plus the picture of a little hand (also in presentation identical to the one used in the picture versions) that moved towards the button; a shadow effect indicated the pressing of the button and the picture of the little hand moved away again. Then the next picture moved in, the hand moved towards the button, etc.

The spatial information informed the subjects where approximately to look for the buttons to be pressed; no exact information was given, as this would include operational information. In all versions the Bieger & Glock taxonomy is applied on a more abstract level than Bieger & Glock did at the paragraph or total picture level (for justification, see the previous chapter).

The most complex tasks required procedures consisting of seven actions. It was immediately clear that it is impossible to see and remember an animation presenting all these seven actions and then perform these actions on the telephone. Pretests showed that subjects did best when performing steps of three actions. Steps of two actions forced them to look at the screen more often, which caused that subjects needed more time. Steps of four actions also took more time, in this case probably because of memory problems. After such a chunk with three actions of a procedure, the subject had to press the OK button on the screen (with a computer mouse) to see the next chunk of the procedure.

For each version 10 subjects were asked to perform 13 tasks with an increasing level of difficulty (table 8.1). The tasks and the instruction were presented on screen. Subjects were instructed in advance how they could navigate from one task and instruction on the screen to the next. The computer measured the times used per subject per task.

Table 8.1. Tasks [translated from Dutch]

1	Change the bell- tone from 99 to 79 (melody 7, speed 9).
2	Change the pin- code from 123 to 321.
3	Set up a dial- block.
4	Program a call forwarding in case of absence from this phone to phone 8972.
5	Switch off this call forwarding in case of absence.
6	Program number 0015-2788971 on the marked key.
7	Call 0015-2788971 using this key.
8	Switch off the dial block.
9	Change the pin code from 321 to 123.
10	Switch on the answering machine.
11	Check if the answering machine is on.
12	Switch off the answering machine.
13	Change the bell tone from 79 to 99 (melody 9 and speed 9).

Tasks 8, 9, 12 and 13 were nearly identical with tasks 1, 2, 3 and 10: switching on versus switching off several functions. This would enable drawing conclusions about the short-term learning curve for the various versions. Tasks were to be performed on two identical telephones.

After each of the subjects had successfully completed the 13 tasks, some questions were asked. The relevant question for this study was: 'Describe the basic system for storing features in this telephone system'. This question was asked to measure the understanding of the system. Results were noted on a three-point scale from:
1 = no understanding to 3 = total understanding.

Almost all of the 60 subjects were students from the Delft University of Technology; most of them were 2nd and 3rd year students from the faculty of Industrial Design Engineering, about equally divided between male and female (slightly more males). This selection of subjects was done for practical reasons; it is obviously not a good representation of the general public: students at a university of technology are likely to be more intelligent than average, be more interested in technology and technical products, have more interest in and experience with complex technological products, etc. Moreover, they are of a limited age group, all between 20 and 25. After completing the test, each student received ƒ 10,-. The subjects were scheduled per half hour.

8.3 Results

The results were measured in seconds and are presented in table 8.2 and figures 8.7- 8.11. Detailed results and analysis are presented in Westendorp (1996b); here only totals and means per version plus standard deviation are presented.

Table 8.2. Mean times per subject in seconds and standard deviation. N=10 per version.

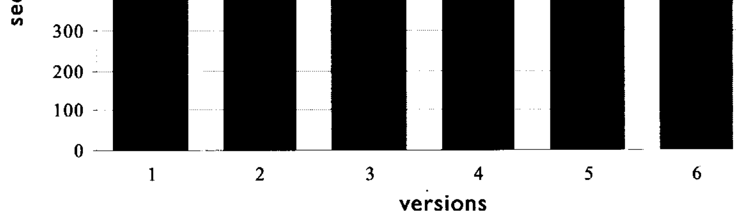
	version 1	version 2	version 3	version 4	version 5	version 6
Mean	422	390	496	461	872	666
Standard deviation	45	33	64	87	114	68

version 1 Operational information in text; spational information (separate) in pictures
version 2 Operational information in text; no spational information
version 3 Operational information in pictures; spational information (separate) in pictures
version 4 Operational information in pictures; no spational information
version 5 Operational information in animation; spational information (separate) in pictures
version 6 Operational information in pictures; no spational information

8.4. Conclusions

The version with only operational information in text (no spatial information) produced the shortest performance times (figure 8.7). Presenting operational information in pictures or with animation proved inefficient. Adding spatial information also seems to be inefficient for all versions, although the difference is not significant.

Figure 8.7
Mean total times per version (in seconds).



8.4.1. Text, pictures or animation

The results show that subjects needed more time to complete all tasks on the telephone system successfully when the instructions were presented in animations than when presented in text or pictures. Also, subjects needed more time when the instructions were presented in pictures than when presented in text (figure 8.8). This is both true for the versions with the added spatial information and the versions without spatial information. An ANOVA test of the results indicates that the differences between the results for the three types of media are significant ($F = 81.7529$, $df=57$, $p < .001$).

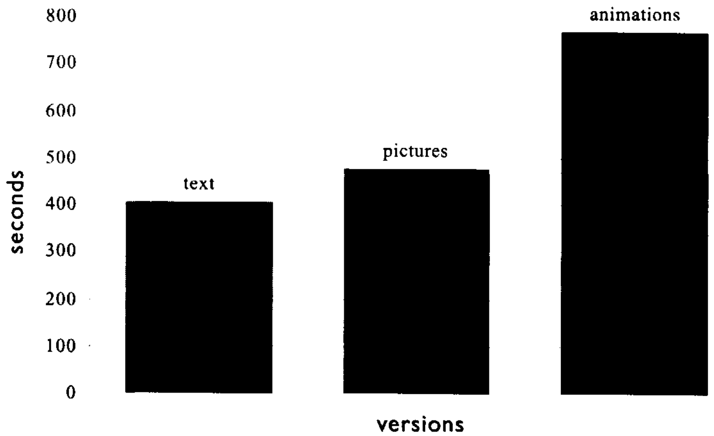


Figure 8.8. Mean total times for the 2 text versions, the 2 picture versions and the 2 animation versions (in seconds).

The overall conclusion is that subjects using the versions with the operational information presented in animation needed substantially more time than subjects using either the versions with this information presented in text or in still pictures. This implies that the original hypothesis must be rejected. More specifically, it implies that subjects using instructions with operational information presented in text (versions 1 and 2) completed the tasks 18% faster than subjects using instructions in pictures (versions 3 and 4) (significance: $F = 4.439$, $df=38$, $p = .042$). Presentation of the operational information in text helped subjects to perform the tasks 47% faster than when this information was presented in animation (significance: $F = 26.134$, $p < .001$). Presentation of the operational information in pictures helped subjects to perform the tasks 38% faster than when this information was presented in animation ($F = 9.982$, $p = .003$). Presenting operational information in text lead to significantly faster successful completion of the tasks than presenting this information in pictures, and to much faster succesful completion of the tasks than presenting this type of information in animation.

On the basis of these findings the conclusion seems justified that, for presenting operational information, text is the most efficient medium and animation is the least efficient. It should be stressed, though, that this conclusion is restricted to applying the information immediately. It is unclear what the effect of these three media is on learning (in this case, memory). For instance, animation might lead to a much faster successful replication of the tasks one week later. This is subject to further research (chapter 9). Also, it should be taken into account that the subjects for this test were all students; results could be different with other subjects, such as people who are older or have less education or less experience with computers and modern electronic products.

From a theoretical point of view it is interesting to note that these figures show that congruency between encoding and reproduction media did not lead to better performance[65]. This contradicts what Mandl & Levin (1989, p. 12) reported when discussing research by Seel & Strittmatter: 'congruency between the encoding and the reproduction media leads to better performance'. In the present study, applying the most abstract and incongruent medium (text) leads to the best performance times. The more realistic medium (pictures) increases performance times and the most realistic medium (animated pictures) leads to even longer times to complete all tasks successfully. Therefore, from this study it seems that encoding plus decoding the incongruent medium (text) is more efficient than understanding, reproducing and applying the more congruent (operational) information from pictures and animation. It is difficult to find an explanation for these results, because they contradict the conclusions as reported by Mandl & Levin. Also, one would expect that subjects would much faster find the elements on the machine that have to be

---

[65] 'Congruency' here means that the instructions look like the elements on the apparatus on which the instructions have to be performed. Text has basically no congruency, while pictures can have a high congruency (which is the case in this study because the pictures are almost identical in shape, size and color) to the elements on the telephone. Animation has the highest congruency, because not only the pictures are congruent with the elements on the telephone, but also the movements look like those that the user has to make. In other words, text is most abstract, animations are most realistic (see chapter 5.5).

pressed successively if they have been presented in the help system as realistically as possible. If text is the instructional medium, users first have to make some kind of small conversion in their minds from the text to the visual element for every button they have to press. One would expect this to cost more 'computing power' of the mind and thus more time. The opposite seems true. A first, tentative explanation could be that in this test the differences are too small: the most crucial element of the pictures was not the button itself, but the symbols on or near them, and these symbols often consisted of text (e.g., the word 'program'). However, sometimes the symbol consisted of an icon or other pictorial symbol (e.g., a small loudspeaker). A very tentative explanation for the results from the study presented here could be that, although the pictures were very simple (just a realistic representation of the buttons that had to be pressed), they may have been somewhat more of a cognitive load for the subject than text. This would be consistent with the view (Schnotz, 1991 cited in Strothotte & Strothotte, 1997, p. 52) that 'all pictures have to be 'read' and 'the capabilities for this process have to be learnt' - at the cost of extra 'computing power'. Once we have learned the rules (language, text), applying the more abstract system could indeed be more efficient than using a new system (= new pictures for every new situation) for every new situation.

8.4.2. Pictures for communicating spatial information

Subjects using the three versions with spatial information needed (15%) more time to complete the tasks successfully than the subjects using versions without spatial information (figure 8.9). The differences are statistically not significant ($F = 3.845$, $p = .054$), yet they may indicate that the use of pictures for spatial information does not help people to perform procedural tasks on 2-D interfaces. For these kinds of products, adding separate spatial information in pictures seems counterproductive. Especially in the animation version, the more information users get, the more difficulty they have in performing the tasks. It seems that for the small control panels (e.g., for car stereo, VCRs, telephones and programs on computer screens) it is faster to read what you have to do and then find the buttons to push yourself, than read what you have to do plus read where to find the buttons and then apply this information.

An explanation could be the following. The cognitive load is bigger when subjects have to remember two different types of information for several steps, or perhaps even for each step. This may indicate that it is useless or even counterproductive to add pictures in instructions for 2D products to help users find what they are looking for. It is still unclear to what extent a conclusion like this can be drawn.

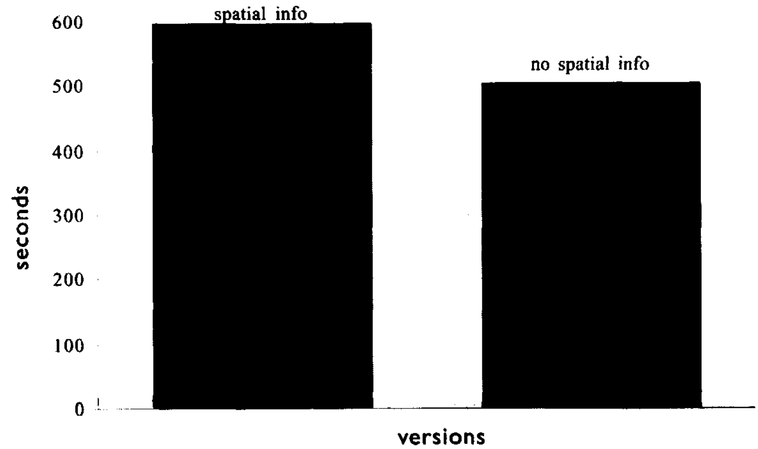


Figure 8.9. Mean total times for versions with spatial information (in seconds).

8.4.3. Understanding of the system

Understanding of the telephone system is very high for all versions (figure 8.10). This is especially interesting because no contextual information was given at all (Bieger & Glock found contextual information to be essential). Subjects were only informed about the tasks and about what buttons to press (and where to find the right buttons in versions 2, 4 and 6). Yet, after performing the tasks, almost all subjects had a pretty good idea of the system (between 2.5 and 3 on a 3-point scale). Performing new (comparable) tasks on this telephone would probably be easy for them.

This could lead to the following conclusions. The instructions given in this test only consisted of procedural instructions; no overall information was presented that explained the telephone system and no attempt was made whatsoever to explain the basics or regularities behind all features. The subjects had to learn this by doing. This 'learning-by-doing' (Gagne & Briggs, 1979; Carroll, 1990) appeared to be very efficient for user instructions. Yet in many user manuals, all features of the product and the results of all kinds of actions are explained first. Many instructions for use seem to be designed and written by 'teachers' who really want to teach the users something by informing them. But it is questionable whether users want this information. Their main goal is to have the machine do what they bought it for. To do so, they need (procedural) instructions much more than basic or overall information. This is also one of the most important points of Carroll's minimal manual theory, which has dominated research into instructions for a long period of time (Brockmann, 1998, chapter 9, especially pp. 416-424).

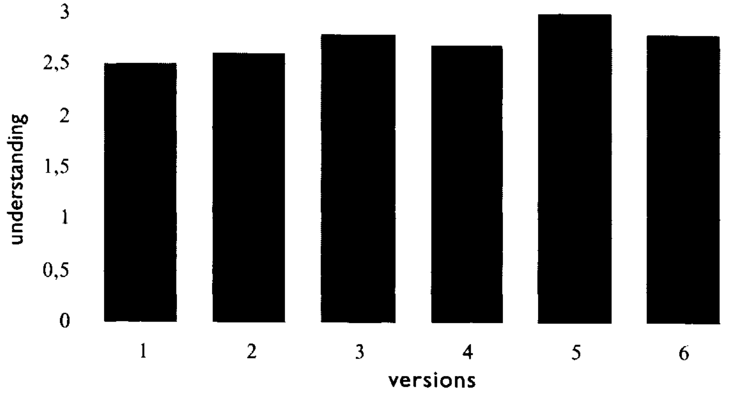


Figure 8.10. Scores for understanding of the system for the six versions.
1 = not (much)
2 = half
3 = (almost) all

Moreover, there are many tasks to perform on modern electronic products and software that users will probably perform only once or very few times (like setting the sound of the bell tone, programming the channels on a VCR, setting the time, installing a mouse or printer driver). It seems inefficient to expect users to first try to understand the basics of the machine and only then start to operate it; users learn this by simply performing the necessary actions to successfully activate all features. For many tasks it may be better to suppose - and take as a starting point for designing the instructions - that users learn by doing and do not want to learn by reading first. The results presented here indicate that they learn very well to extract the more abstract rules on which the system is based.

For user interface designers this reasoning implies that it may be more important that all features are activated and de-activated according to the same principles as much as possible. Designers and writers of user manuals and on-line help systems should rely much more on the learning-by-doing principle and only inform users what to do. In user manuals 'teachers' tend to be dysfunctional. People learn much faster and better from the things they discover themselves then from what others tell them (Gagne & Briggs, 1979). Of course it must be remembered that this argumentation is not based on a study contrasting instructions with and without contextual information, but only on the fact that subjects in this test had developed a good understanding of the system just by doing.

Subjects using the animation-based instructions (versions 5 & 6) had developed a better understanding of the system than subjects using the picture-based versions (4 & 5); subjects using picture-based versions had developed a better understanding than subjects using the text-based versions (1 & 2). The most obvious explanation for this is that this has to do with the time users needed to perform all tasks successfully: the more time they needed, the better their understanding of the system became.

No conclusion can be drawn concerning the effect on understanding about the addition of spatial information: in the text versions this led to a slightly better understanding; in the picture and animation versions it led to lesser understanding.

8.4.4. Learning speed

Subjects always needed more time when using the animated instructions, but they learned much faster when animation was applied (figure 8.11). Performance times dropped sharply for the nearly identical tasks at the end of the tests, but were still a little bit above those of the text and picture versions. Performance times for the nearly identical tasks at the beginning and at the end of the tests dropped 36% for the text versions, 53% for the picture versions and 68% for the animation versions. The tasks at the beginning of the tests took the subjects using the animation versions almost three times as long as the subjects using the text versions, but for the identical tasks at the end of the tests they needed only about 30% more time. Perhaps animation is more useful if still longer series of (identical) tasks have to be performed or if the information has to be learned for repetitive application over a longer period of time.

It seems that, in spite of the communication possibilities provided by modern technology, offering only operational information presented in text to users of feature-rich products is the most efficient way to help them. Perhaps in situations where users want to explicitly learn something (e.g., at school), animation is a more efficient medium. When users simply want to perform certain tasks, the best communication is as minimal as possible, and in text only.

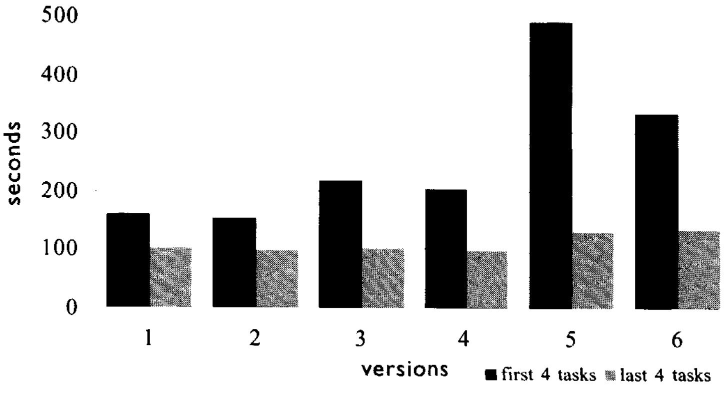


Figure 8.11.
Mean times subjects needed to complete the first 4 tasks and the nearly identical last 4 tasks succesfully (in seconds).

CHAPTER 9

Learning instructions in text, pictures and animation

[This chapter is based on: Westendorp, P.H. (1996c), *Learning efficiency with text, pictures and animation in on-line help*. Journal of technical writing and communication, Vol. 26(4), 401-417.]

9.1. Introduction

User instructions are usually a combination of (1) instructions that users do need to learn (for operations that have to be performed regularly), and (2) instructions that designers cannot and should not expect users to learn (for operations that they have to perform only very occasionally). For instructions that are executed only very occasionally, textual presentation of the operational information seems the best choice. For instructions that should be learned (so that replication of the tasks is easy) or systems that should be understood (so that users can derive other procedures themselves), animation could be an efficient presentation medium (§ 8.4 and figure 8.11).

In the previous chapter, a study was reported that contrasted the relative efficiency of text, pictures and animation for operational information in instructions for the use of a telephone system. Text proved to be the most efficient medium for presenting the operational information; animation appeared to be the least efficient. It was concluded that, for actions that had to be performed immediately, animation was not the best medium in instructions for the use of complex electronic products. But in that test there were also some indications that subjects had learned[66] better from the animation versions. They had improved more noticeably when [1] they had to perform nearly identical tasks again in the same test and [2] they had understood the basic idea of the telephone better (§§ 8.3 and 8.4, figures 8.10 and 8.11).

---

[66] 'Learning' is specified here as the combination of pattern-recognition, abstraction, memory and understanding, manifested in a time-curve of improving performance.

[1] The first four tasks and the last four tasks subjects had to carry out were almost identical (e.g., task one was setting the bell sound from tone 99 to tone 79, and task thirteen was to set it back from 79 to 99). Especially for the first four tasks, subjects using animation versions needed much more time to perform the tasks successfully than subjects using the text and picture versions. For the nearly identical last four tasks, though, the differences were much smaller (§ 8.4 and figure 8.11).

[2] Subjects using the animated instructions appeared to have a better understanding[67] of the telephone programming system. The text versions scored 2.5, the picture versions 2.7 and the animation versions 2.9 on a 3-point scale (§ 8.3 and figure 8.10). Therefore, animation may be not very efficient for direct actions, but could be an efficient medium when remembering and understanding are important aspects. Of course, it must immediately be realized that this variation in understanding correlates with the time subjects needed with the three media: the shortest time with text versions, the longest time with animation versions. The better understanding that subjects using the animation versions appeared to have may be explained just by the fact that they had spent more time.

In the tests mentioned in the previous chapter, the time between the first four tasks and the last four tasks was limited: all tasks were performed within half an hour. The repetition of the nearly identical tasks was done within this half hour, often within less than 15 minutes. So, the time users had to remember the series of actions for these tasks was very limited. This raised the question what the long-term efficiency would be of animation-based instructions in contrast to text-based and picture-based instructions.

9.2. Specification of content and presentation

To be able to draw generalizable conclusions about the relative efficiency of animation compared to other media (text, pictures) for user instructions, we must identify both [1] content and [2] presentation of the information in user instructions in generalizable terms.

---

[67] 'Understanding' was studied by asking subjects to describe the principle for storing function-settings of the telephone-system.

[1] Aspects concerning content have been discussed in the previous chapters, leading to the application of the Bieger & Glock taxonomy of information types. This taxonomy describes the content of the instructions in nine well-defined and identifiable types, which will be applied in this test in the same way as it has been applied in the previous tests and with the same criticism in mind. As a result, conclusions can only concern a specified type of information in the instructions, not the instructions as a whole.

[2] Generalizable conclusions also require a generalizable description of the presentation (or 'style') of the various media. It must somehow be guaranteed that text, pictures and animations are representational for all or for an important category of instructions. A generalizable description of the presentation of the texts, pictures and animations in contrastive tests is reached more or less by (1) the specification and limitation of the types of information that are compared and by (2) the limitation to products with two-dimensional user interfaces.

[2a] Specified categories of instructions in text seem relatively easy to describe in generalizable terms and are relatively easy to control. For instance, in user instructions, the texts conveying the operational information usually consist of short, imperative sentences, mainly consisting of verbs and nouns, without personal pronouns. Spatial information is conveyed by prepositions ('left', 'beneath', 'next to'), and adverbs of place ('backwards', 'pointing-up'). Inventory information is usually presented in nouns ('display', 'program button', 'follow-me button'). Contextual information is presented in prose style, using complete, descriptive sentences, often in the passive voice ('Follow-me' is used to direct incoming calls to another telephone'). This has been a guideline for the design of the textual versions in this and in the previous tests.

[2b] It is the wide variety of possible pictures and the lack of a system to analyze and describe them in particular that pose problems in identifying the information that is communicated by the pictures in generalizable terms. Yet, it may be possible to make a comparison of various media for instruction that leads to generalizable conclusions.

As the tasks for the the subjects of these products had to be performed on two-dimensional products, the pictures could be applied similarly. This makes comparison of text and pictures (and animation) less arbitrary than in cases where three-dimensional pictures (indicating depth) are applied (see § 7.1). The guideline for additional instructional elements, such as arrows (indicating a part or a direction or the direction of a movement) is that these must be conspicuous in relation to the parts of the picture that depict reality. Because the animations tested in these studies used the same pictures as the picture versions, the same guideline applied for these as well.

Separation of types of information and the use of simple sentences and only two-dimensional pictures reduces, but does not eliminate, the problem of comparing a bad text with good pictures or vice versa as described by Wright (1988), Sims-Knight (1992) and others (chapter 6). For the instructions for modern electronic products, we can compare reasonably standardized text with reasonably standardized pictures and with reasonably standardized forms of animation if the tests are limited to two-dimensional user interfaces and instructions.

9.3. Operational and spatial information only

The main question in this test concerns the relative learning-efficiency of text, pictures and animations respectively for specific types of information. It is a follow-up test of the one described in the previous chapter. The same two types of information are involved and for the same reasons. Operational information is included, because this type of information is obviously vital in instructions. Spatial information is included, because Bieger & Glock found spatial information to be essential (chapter 6). Moreover, for both these types of information both text and pictures (and animations based on these pictures) could be relevant media.

If operational information is left out, the instructions may not be understood correctly and efficiently. But it is questionable whether spatial information is really efficient for products with two-dimensional interfaces, because the buttons, icons, symbols, windows and displays are presented on such a small area that the user may find the right button faster by looking at the control panel or screen rather than look in the user manual, remember and then look on the panel or screen (which leads to a higher cognitive load). Therefore, it seems relevant to not only compare the efficiency of text, pictures and animation for operational information, but also to look at the efficiency of adding spatial information. This will always be presented in pictures.

9.4. Method

It is hypothesized that learning from user instructions is more efficient when operational information is presented in animation than in pictures or text, and more efficient in pictures than in text. It is further hypothesized that learning from user instructions is more efficient when spatial information is added (in pictures).

To test these hypotheses, the test that is described in chapter 8, was repeated. The difference is that in this study, the subjects were also asked to perform the same 13 tasks again one week later, using the same instructions, and presented in the same medium as each of the subjects had used the week before.

Therefore, the same six variations of instructions were presented to the new subjects for the completion of a number of tasks to be performed on a telephone system.

1. Operational information in text and spatial information in pictures.
2. Operational information in text and no spatial information.
3. Operational information in pictures and spatial information in pictures.
4. Operational information in pictures and no spatial information.
5. Operational information in animation and spatial information in pictures.
6. Operational information in animation and no spatial information.

For screen-dumps of the versions, see chapter 8).

In this study too, the spatial information informed the subjects where to look approximately for the buttons to be pressed; no exact information was given, as this would include operational information. In all versions the Bieger & Glock taxonomy is applied on a paragraph or total picture level, which is a more abstract level than Bieger & Glock used. For a discussion of the abstractness problem, see chapter 6.

For each version 10 subjects (all students at the Delft University of Technology, mostly 2nd and 3rd year students from the Faculty of Industrial Design Engineering; male and female more or less evenly distributed) were asked to perform 13 tasks on a telephone system. These subjects were not the same as the subjects who participated in the previous test; this test was done one year later. The tasks and the instructions were presented on a computer screen situated next to the telephones. Each screen always showed the task in combination with (part of) the instruction. Subjects were instructed in advance how they could navigate from one task and instruction on the screen to the next. The computer measured the times used per subject per task. The animation was again presented in steps each consisting of three actions. Subjects could replay the animation in part or in its entirety for each task as often as they wanted by one simple mouse-click on a screen button.

The 13 tasks were identical to those of the previous test (table 8.1). They were presented in exactly the same order, which meant an increase in the level of difficulty. The tasks were supposed to be representative of the kinds of tasks users have to perform on electronic consumer and office products. The simplest task took three actions (steps) to perform, the most complex tasks seven actions. The four tasks at the end were nearly identical with the four tasks at the beginning (switching on versus switching off the several functions). In this study too, tasks 1, 2, 3 and 4 were nearly identical to tasks 8, 9, 12 and 13 (switching on versus switching off several functions of the telephone). This was done again to measure the short-time learning curve for the various versions.

After a subject had successfully completed the 13 tasks, the test supervisors asked the following questions and noted the answers on sheets.

1. 'Describe the basic system for storing features in this telephone system'.
   This question was asked to measure the understanding of the system. Results were noted on a three-point scale:
   1 = no understanding; 2 = more or less understood; 3 = complete understanding.
   If a subject had not recognized a pattern, a '1' was noted; if a subject could reproduce one or two elements of the procedure (always start with 'hands free', then always press 'program', etc.), a '2' was noted; if a subject could reproduce the complete abstract procedure (for instance: always 'hands free', then 'program' then a selection of possible functions, then a variable, then 'program' then 'hands free'), a '3' was noted.

2. 'Describe what you have seen on the screens'.
   This question was only asked to subjects using instructions with

separate pictures presenting spatial information. The intention of this question was to find out how much of this additional information was indeed perceived. If a subject could not reproduce any of this information, a '0' was noted; if only the picture was noticed a '1' was noted; if the arrow and/or the red color indicating the relevant area at a certain step was also noticed, a '3' was noted.

All subjects were asked to repeat the same 13 tasks in the same situation and with the same version of the instructions exactly one week later. Also, the same questions were asked again. Subjects received *f* 20,- after completing both tests. If a subject did not show up for the second test, the data for this subject were removed, and a new subject was asked to do the first test followed a week later with the second test.

The times that were measured in the test include both the times subjects needed to consult the instructions and the times they needed to perform the tasks. To measure long-term learning efficiency, the times needed to consult the instructions must be deducted from the overall times, so that the times needed just to operate the telephone system can be compared. This is important since it was clear from the previous test that especially subjects using the animation versions needed much more time to consult the instructions. One reason for this was that they had to press the 'Next' button after every three steps of a procedure. Because procedures sometimes consisted of 7 steps, some procedures needed 3 screens; subjects not only had to press the 'Next' button to proceed to the next screen but occasionally also the 'Back' button to return to a previous screen or the 'Home' button to return to the start screen of this procedure.

To measure the mean times needed to read (text), see (pictures) or page through and see (animations) each version of the instructions, five persons who knew both the instructions and the telephone system very well (two test supervisors and three designers of the instructions, none of whom were subjects in the test) read aloud all the tasks and looked at all the instructions on the screens, but they did not perform the tasks on the telephones. They times they needed to do so were measured.

The software that was used to design the instructions (MacroMedia Director 3.0) also registered the times the subjects and the five persons that did the pretest needed. These times were registered in tics ($^1/_{60}$ sec.); data were transferred to Excel, which was used to recalculate in seconds and diagram the results. Significance (on a 5% level) of the data was checked with paired means and ANOVA tests, using SPSS 6.0.

9.5. Results

The pretest involved five specialists to read aloud the tasks and then read the text versions or look at all the pictures in the picture versions and leaf through and scan all the pictures in the animation versions. Each of these specialists executed all of these versions in varying order. It appeared they needed 20 seconds on average for the text versions, 20 seconds for the picture versions and 186 seconds for the animation versions.

The results are presented as means with the standard deviations for the ten subjects for each version in figures 9.1-9.6. For detailed results and further analysis see Westendorp, 1997b; here only mean total times and standard deviations for each version are presented.

version 1 Operational information in text, spatial information in pictures.
version 2 Operational information in text, no spatial information.
version 3 Operational information in pictures, spatial information in pictures.
version 4 Operational information in pictures, no spatial information.
version 5 Operational information in animation,; spatial information in pictures.
version 6 Operational information in animation, no spatial information.

Table 9.1. Mean total times for correct performance of all tasks (in seconds) and standard deviations. Test 1. N=10 per version.

	version 1	version 2	version 3	version 4	version 5	version 6
Mean	444	402	553	457	784	759
Standaard deviation	18	39	27	35	35	28

Table 9.2. Mean total times for correct performance of all tasks (in seconds) and standard deviations. Test 2. N=10 per version.

	version 1	version 2	version 3	version 4	version 5	version 6
Mean	356	315	334	289	354	367
Standaard deviation	14	30	17	18	27	18

Figure 9.1 presents an overview of the mean total time that subjects needed to complete the 13 tasks successfully. The left columns show the results of the first test; the columns on the right show the results of the second test (one week later).

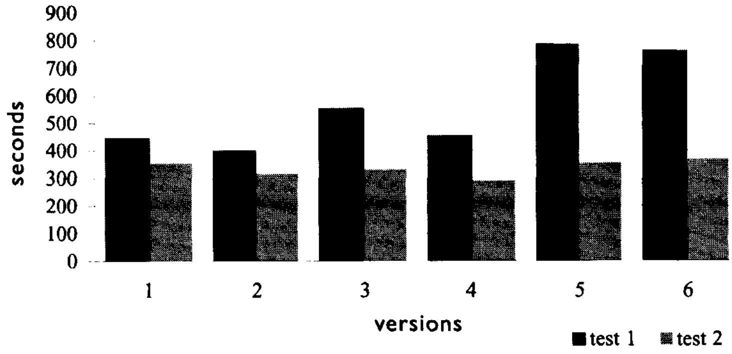


Figure 9.1. Mean total times per version (in seconds). Results of the tasks of the first week (test 1) compared with the results of the same tasks done by the same subjects, using the same versions of the instructions one week later (test 2).

Figure 9.2 shows the mean times for the text, and animation versions. The left columns in this chart present the data from the first test and the right columns present the data from the second test.

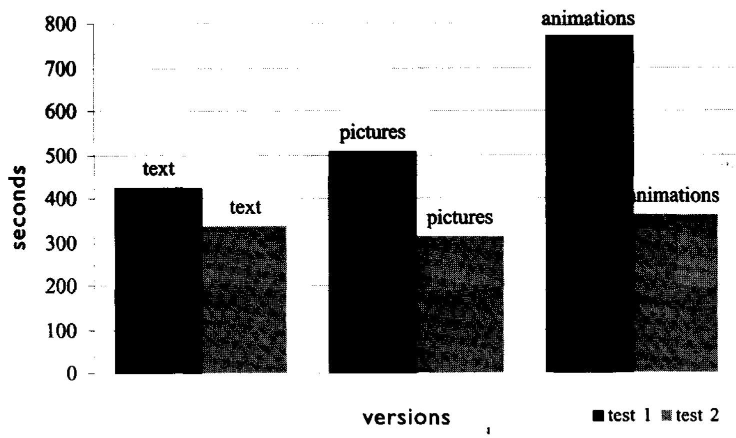


Figure 9.2. Mean times for completion of the tasks (in seconds) for text versions, picture versions and animation versions.

In the text versions and in the picture versions the instructions for each of the 13 tasks could be presented on a single screen. Five specialists, with a good knowledge of the telephone system and the instructions, needed 20 seconds on average to read these instructions. In the animation versions the instructions were presented in steps consisting of three actions. This resulted in the presentation of the information in up to three screens per procedure. The three different presentations led to quite different times that 5 subjects needed to consult all instructions for all 13 tasks. The data presented here are means for 5 subjects who knew the telephone system well. The data are also averaged for the two versions for each medium (two text versions, two picture versions, two animation versions).

Table 9.3.
Mean times needed to read all 13 tasks and consult all instructions for each of these 13 tasks (in seconds). N=5.

medium type	time
textual instructions	20
pictorial instructions	20
animated instructions	186

Figure 9.3 shows the mean total times minus the average times that five persons needed just to read/see/page through the instructions (table 9.3). As a result, the calculated times needed to perform the tasks without the times needed to consult the instructions are presented. Compare with figure 9.1: for text versions (1 & 2) and picture versions (3 & 4) 20 seconds were deducted; for animation versions (5 & 6) 186 seconds were deducted.

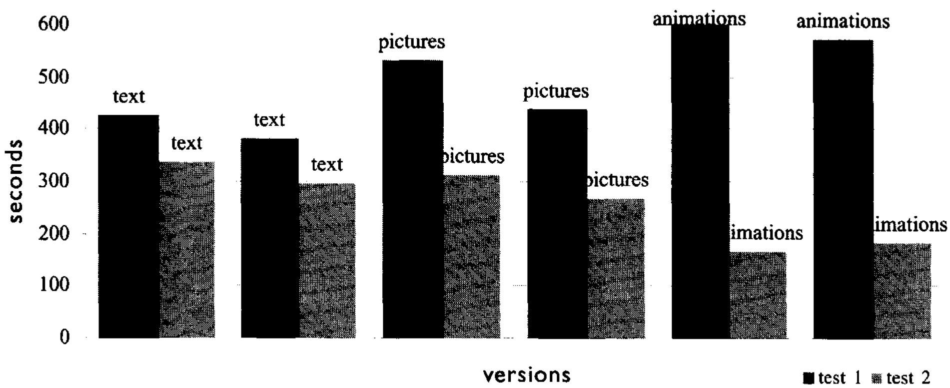


Figure 9.3.
Corrected mean total times for completion of the tasks for the 6 versions (in seconds). The times needed by an expert to read/see all instructions deducted from the total times (presented in figure 9.1).

Figure 9.4 shows the mean times of the versions with and without the spatial information in separate pictures. The left columns show the data derived from the three versions with spatial information in separate pictures; the right columns show the data derived from the three versions without this information.

Figure 9.4.
Mean times for the versions with spatial information and the versions without spatial information in test 1 and test 2 (one week later) (in seconds).

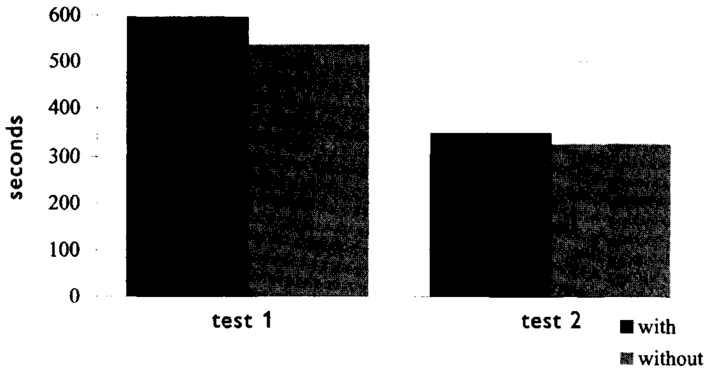


Figure 9.5 shows how much of the spatial information in separate pictures was perceived which was only presented in versions 1, 3 and 5.

Figure 9.5.
Scores for the reported perception of separate spatial information.
0 = nothing;
1 = one element;
2 = two elements;
3 = all.

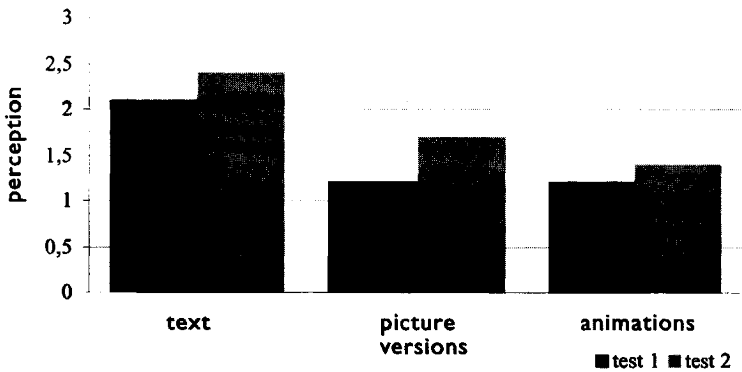


Figure 9.6 shows the understanding of the system for each version in both weeks on a three-point scale. The left columns present the data from the first test; the right columns present the data from the second test.

Figure 9.6.
Scores for understanding of the system for the six versions in test 1 and test 2 (one week later).
1 = not much;
2 = half;
3 = (almost) all.

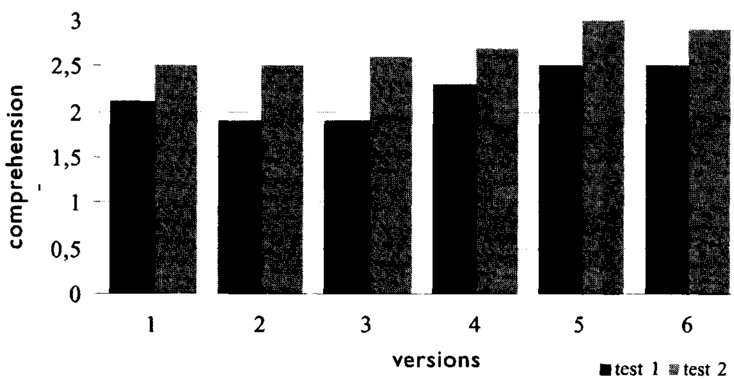


Figure 9.7 shows the results of the first four and last four tasks in test 1. These tasks are nearly identical (switching on and switching off the same features of the telephone).

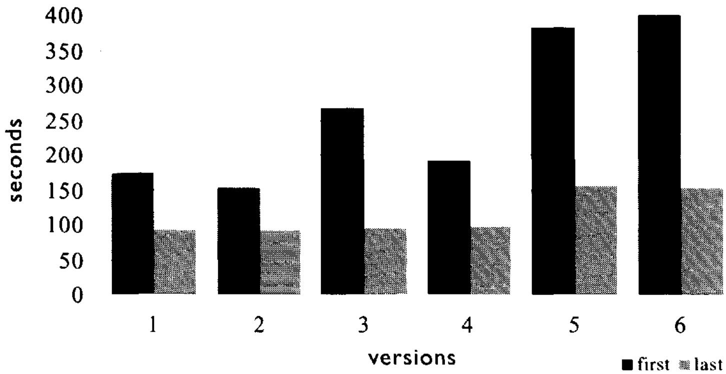


Figure 9.7.
Mean total times to complete the first four and last four nearly identical tasks successfully in test 1 (in seconds).

Figure 9.8 shows the results of the first four and last four tasks in test 2. These tasks are nearly identical (switching on and switching off the same features of the telephone).

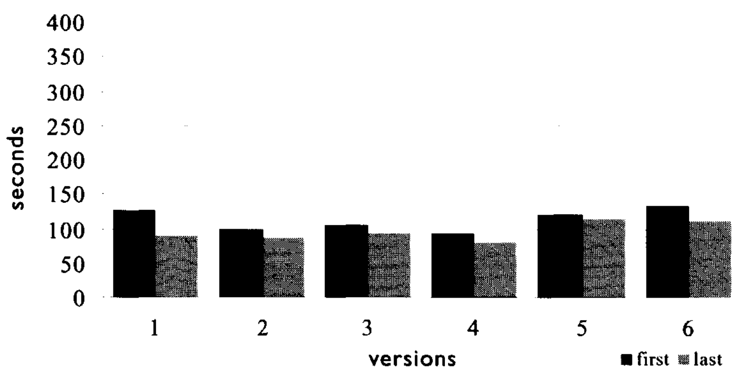


Figure 9.8.
Mean total times to complete the first four and last four nearly identical tasks successfully in test 2 one week later) (in seconds).

9.6. Conclusions

Just as in the previous test (chapter 8), presentation of the operational information in text, without adding the pictures with spatial information, appears to be most efficient (mean total time subjects needed: 402 seconds; see figure 9.1, test 1, version 2). Likewise, presenting the operational information in animation, and including separate pictures for the spatial information, appears to be the least efficient (mean total time subjects needed: 784 seconds; see figure 9.1, test 1, version 5). The versions with the operational information in pictures score in-between. The results are close to the results of the test reported in the previous chapter and support the conclusions based on those tests. An ANOVA test of the results indicates that the differences between the results for the three types of media are significant ($F = 288.047$, $df = 59$, $p < 0.001$.

In the second run of this test, one week later, the differences between the six versions are much smaller (between 289 seconds and 367 seconds; see figure 9.1, right columns), but the differences are still significant ($F = 18.921$, $df = 59$, $p < 0.001$). Of course, one week later, the same subjects performing the same tasks with identical instructions performed better. The results show that there is a significant reduction of the times subjects needed to complete all tasks successfully: all differences with the results from the test in the first week are significant (all t-tests for paired samples have a significance < 0.001). This difference is largest for the animation versions and smallest for the text versions, so that in the second test results for all versions are much closer.

9.6.1 Text, pictures or animation for operational information

Just as in the previous test (chapter 8), text is significantly more efficient for presenting the operational information than pictures ($t=5.34$, $df=38$, $p < 0.001$) and pictures are significantly more efficient than animation ($t=17.82$, $df=38$, $p < 0.001$) (figure 9.2).

In test run 2 , one week later, applying the operational information still took significantly more time in the picture versions than in the text versions ($t=-2.57$, $df=38$, $p < 0.014$) and more time in the animation versions than in the picture versions ($t=5.90$, $df=38$, $p < 0.001$), but the differences are small (336; 311 and 361 seconds on average per subject respectively). Subjects have learned to work with the animations and apply the information more efficiently than with the text and picture versions. Yet, if there would be no further improvement, it would still be better to present the information in text, as long as subjects still need to work with the instructions.

9.6.2 Operation times only

To read the tasks and the instructions in the text versions takes about 20 seconds; to read the tasks and look at all the pictures in the picture versions also takes about 20 seconds. To read the tasks and proceed through the animations in the animation versions takes much longer: about 186 seconds on average per subject (table 9.3). This difference blurs a comparison of the learning effect, because 'operating' the animated instructions takes more time than consulting the instructions in texts or pictures. These times should not be included in a comparison to make specific conclusions about the learning effect. If we subtract these mean times needed to read/see/go through the instructions in the six versions, we obtain the calculated times needed just to perform the tasks (figure 9.3).

In the first test, version 2, with operational information in text and no spatial information, is still the most efficient one and version 5, with operational information in animation plus spatial information in pictures, is still the least efficient ($p < .001$). Overall, the data show the same trend: texts most efficient, animations least efficient. However, the differences are much smaller.

The effect of deducting the times to read/see/go through the information is especially interesting if we look at the times needed in the second test, done one week later. Then we see that the animation versions appear to be more efficient than the text-versions and picture versions (168 and 181 seconds for the animations; 337 and 295 seconds for the text versions and 315 and 269 for the picture versions). It seems that subjects working with the animation versions in the second week knew quite well where to find the right buttons and in what order to press them. Many subjects pressed the buttons on the telephones while the animations were playing - a clear indication that they had memorized the instructions. In this sense, they had learned more from the animation versions than from the text and picture versions.

9.6.3. With or without pictures for spatial information

Adding separate spatial information in pictures again proved inefficient (figure 9.4). In fact, subjects seem to perform less well if they get this information (with spatial information 594 seconds;

without 539 seconds). The differences are small and not significant ($t=-1.36$, $df=58$, $p < 0.179$), but it is clear that presenting this information is not very helpful and possibly counter-productive. This does not change when the tasks have to be performed again one week later. Subjects using the versions with spatial information performed a little (but significantly: $t=-2.96$, $df=58$, $p < 0.004$) less well than subjects who did not get spatial information (348 versus 324 seconds). So, also for repetition of the tasks, there seems to be no use in presenting the separate spatial information in instructions for products with two-dimensional user interfaces. At least, it does not seem efficient to add the spatial information in separate pictures.

Figure 9.5 suggests that subjects using the text versions have seen more of the additional spatial information than subjects using the picture versions. Subjects using the animation versions could reproduce the least amount of the spatial information. This did not change in the second test. Perhaps the cognitive load is bigger when people have to work with animated instructions, so that subjects have less attention for the additional information.

9.6.4 Understanding of the system

Just as in the previous tests, all subjects seem to have learned very well the basic system for storing features in this telephone (figure 9.5). The text-only version was the least successful, but the differences with the most successful versions (the animation versions) are not very big (between 1.9 and 2.5 on a 3-point scale). After applying the 13 tasks again, one week later, all versions score high (between 2.5 and 3) with all versions. Subjects did learn better from the animation versions, but on average they did pretty well with the text versions and the picture versions too. Subjects did not receive any general information about the basic system for storing features; they seem to have learned very well just by doing. The results of both test runs show the same trend as those from the previous test and support the conclusions presented in chapter 8.

9.6.5. Learning efficiency

Figures 9.7 and 9.8 show the short-term learning speed results. The data in figure 9.7 show the same trend as the data in the previous test (figure 8.11). The short-term learning speed is least for the text versions and greatest for the animation versions. The differences between the three media are rather small for the last four tasks, but animation still takes twice as long.

When subjects had to repeat the tasks one week later, the differences between the text versions, the picture versions and the animation versions are much smaller, as can be seen from figure 9.8. In the previous study and in the first test of this study, the first four tasks took subjects using the animation versions on average almost three times as long as subjects using the text versions. The second run of this test shows the same trend as the first run: small differences between the three versions, animation being slightly less efficient than text and pictures. It is remarkable that there is no further improvement for the animation versions in the second run of this test. It seems that the learning curve has already leveled off (subjects had performed the four tasks four times: at the beginning and at the end of each test run). In this test, the picture versions seemed to be the most efficient, but the differences are very small.

9.6.6. Overall conclusions

In general, text appears to be the most efficient medium for presenting operational information in instructions, while animation appears to be most efficient if the instructions have to be memorized, that is: if they do not have to be used anymore for further use of the product. Moreover, subjects using the animated instructions had understood the basic programming principle of the telephones better than subjects using the text versions or picture versions. Short-term memory did not improve very much in the second test run. The differences between textual, pictorial and animated instructions became very small when subjects had to perform the same 4 tasks at the end of the second test run. It was the fourth time they had performed these tasks (twice in the first week and twice in the second week). The learning curve flattens. Subjects had learned the tricks and the systems.

CHAPTER 10

Guided systems versus interactive on-line help

[This chapter is based on: Westendorp, P.H. (1997a), *Guided systems versus interactive on-line help*. In: Morris, A., Houser, R. & De Loach, S., Cross-roads in computer communication. IEEE/IPCC/ACM conference proceedings. IEEE Publications, Salt Lake City, pp. 1-8.]

10.1. Introduction

Efficiency, effectiveness and attractiveness of user instructions depend on many aspects (chapter 3). Previous chapters have focused on one of these: the selection of media (text, pictures, animation) for user instructions. This chapter will focus on the possibility to make some parts of the instructions in on-line help systems interactive.
'Interactive' means that the user does not have to enter the instruction that is presented on a screen or display into the machine that just presented it, but may directly activate the help system to let the machine perform the action, or part of the action (§ 3.1.3, especially figures 3.4 and 3.5). Help systems with these possibilities are called 'interactive' because users may act directly from within the help mode of the product, rather than just read the information, switch to the standard mode and then enter the instructions into the machine. In such an interactive system the machine may immediately perform the whole task about which the user asked information or lead him or her into a dialogue in which the user, step-by-step, fills in specific data that the machine cannot know, for instance, a number to indicate loudness or a preferred button to store a radio station.

Interactive on-line help systems can be applied in combination with both basic types of user interfaces: guided systems (1) and command-line interfaces (2) (Shneidermann, 1998).
(1) In a guided system (GS) the user is guided towards his or her goal. This can be done in various ways.

- By selecting meaningful instructions from (hierarchical) menus.
- By pressing a number of buttons (icons). If these buttons are on screen, this is called direct manipulation.
- By filling in forms.

In many software programs and products these interaction styles are combined.

(2) In a command-line interface (CLI), the user has to enter codes. These codes are more or less arbitrary, or at least nonintuitive, and often difficult to deduce. The user has to find them by consulting the user manual or the on-line help system. Well-known examples are the DOS and UNIX operating systems, but more generally known are the codes (usually numbers) that have to be entered into VCRs (channel 8), cameras ($^1/_{125}$, f5.6) and telephones (*1121 or ##31#).

Comparison of GS and CLI

Guided systems are usually presented with a complete graphical user interface (GUI) and have generally been accepted as the most user-friendly, especially for non-experts (which means everybody who does not regularly operate a certain feature of the product). On many products a guided system may be hard to realize, because the display is too small; this is, for instance, the case with telephones, cameras and portable radios[68]. For these products a CLI is sometimes the only alternative. A CLI is generally considered to be more efficient once the user has learned the codes by heart, but far less efficient if the user does not know what to do. This is not only the case when users are confronted with a product for the first time, but also every time they want to activate a function that is not used very often, such as programming a new station on the audio set or changing the bell tone of a telephone. That is why CLIs may continue to be a problem, even for experienced users of the basic functions.

10.2. Problem

In a CLI, users have to rely on the help system more than in guided systems. They have to find the (more or less) arbitrary codes that have to be entered; in guided systems, all relevant alternatives are presented and the user just has to select the preferred one. Yet, CLIs may offer a more efficient alternative to guided system-based interfaces for novice users of the features if the help system is fast, intuitive and intelligent. Interactivity of the help system may contribute to this.

---

[68] In recent years, bigger displays are getting more common, even on small products, and some producers have managed to present guided systems on fairly small displays (for instance, the latest generations of mobile phones).

At the moment many products still have to be operated by consulting a separate user manual. This means that a user who wants to activate a feature has to perform a number of actions:

1. define the problem;
2. find where exactly the instruction is presented;
3. understand the instruction;
4. (sometimes) deduce the right actions;
5. memorize the actions;
6. apply the actions.

A detailed example was discussed in table 4.1.
Many problems in interacting with products with a CLI are caused by steps 5 (memorizing the actions) and 6 (applying the actions). This study will concentrate on these two steps.

Memorizing the actions (Step 5)

Usually users have to switch between the information in the user manual or on the display and the buttons on the product (chapter 4). They must learn the instruction (which consists of one or more individual steps and maybe some additional information) by heart, remember it, move the eyes to the buttons, find the right buttons, press them and relocate the right place in the user instructions for the next step or for checking the feedback. Notice that users also have to keep in mind the location of the information in the user manual when remembering the action and finding the buttons. This mini-procedure has to be repeated for every detailed step in an instruction procedure.

During this switching back and forth between the instruction and the product, users sometimes not only have to memorize and apply information, but also process it further. That is the case in steps (action sequences) where users should not press a specifically defined button, but a button that can only be indicated, for instance with an example. This is, for instance, the case when a user has to decide which button (number) to program for a specific television station. In the instructions for this procedure, only an example can be given, because it is up to the user to choose the number. That user reads the example with a specific number, then has to recognize that it represents a rule, and then has to deduce from that rule a specific, preferred number. It is a very low-level abstraction and seemingly simple rule, but the user has to do this many times and it is only a very small part of the complete instruction procedure. This processing during the performance of instruction tasks consisting of various steps increases the difficulty for the users. It adds a little task on a different level of information processing.

Rasmussen (1986) distinguishes three levels of information processing: skill based, rule based and knowledge based (figure 10.1). 'Skill based' means the user applies the information more or less automatically; 'rule based' means he has to deduce information from a certain well-known rule, and 'knowledge based' means the user has to infer the exact what-to-do information him or herself. In user instructions, users may have to combine all three different levels of operation: skill based, rule based and maybe also knowledge based. Many tasks involve just pressing a key that is mentioned or reading something from a display (skill based). Users may also have to apply a procedure that was mentioned before ('repeat steps 4 to 7') (rule based) or find out the rule themselves, for instance, if only an example or the grammatical rule is mentioned (knowledge based).

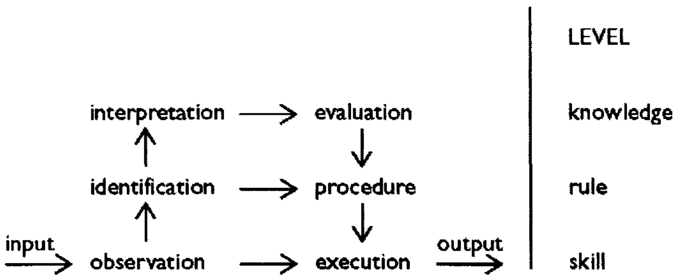


Figure 10.1. Rasmussen's model of information processing on three levels.

Part of this repetitive process of perception and cognition may be done during the switch from information to application (as has been discussed in chapter 4). A user who has to do so many actions (reading, remembering, processing the information, finding the buttons and switching back to the right place in the user instructions) may easily forget or deform the information. In user tests, it is striking how often test supervisors can see subjects read an instruction and do something totally different from what the instruction states.

Applying the actions (Step 6)

Then the user has to apply the instructions, which means pressing or selecting buttons or entering codes. In this phase too, mistakes can easily be made. Usually the user simply enters the information that is presented in the user manual or on the screen or display. Especially when the instructions are presented on a screen or display, this seems a peculiar thing to do: to enter codes that the machine obviously already knows. For example, for changing the bell tone of a telephone, the user has to perform the following 9 steps.

1. lift the handset or press 'hands free'
2. press a program button
3. press #
4. choose a number between 1 and 9 for the volume
5. press #
6. choose a number between 1 and 9 for the melody
7. press #
8. press the program button
9. put down the handset or press 'hands free'

Now, imagine that these steps are presented in the on-line help function on a display. In that case, to change the bell tone, the only things the machine does not yet know are the variables that the user has to choose: the numbers for volume and melody. The other actions are already 'known' to the telephone and, therefore, should not have to be entered from the machine into the machine. Some recent systems software and applications indeed offer help that can be applied immediately, without the user typing what is already on the screen. The first software program to present such interactivity is WordPerfect for Apple Macintosh computers, release 1.0. It had an 'Execute' button, which allowed the user to let the computer perform the selected task (figure 10.2).

This system was probably not a great success, because it was not reused in the next release (2.0) of WordPerfect for Macintosh (1990). In 1995 Microsoft re-introduced interactivity. In the Help system of Windows 95 some examples can be found of buttons in the help system that prompt wizards that help the user perform the task. In that

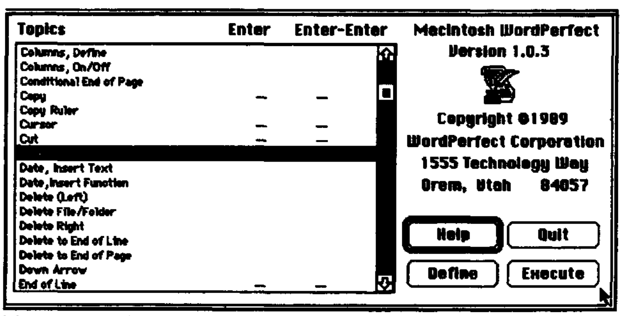


Figure 10.2. Help screen of WordPerfect for Macintosh; note the 'Execute'-button.

Figure 10.2. Help screen of WordPerfect for Macintosh version 1.03 of 1989; note the 'Execute' button.

year WordPerfect Presentations for Windows also had some of this type of agent-directed interactivity. In Office 97, Office 98 (for Apple Macintosh) and Office 2000, this type of interactivity was further developed. More developments on this aspect can be expected from Microsoft (Hoek & Kaufman, 2000) and from Apple Computer (Knabe, 2000 and Arcellana, 2000)[69].

For small products like telephones, the information the user needs is still often presented in the user manual, because displays are still too small to present the help information. If there is any on-line help in these systems, this information in many cases disappears as soon as the user switches back from the help status to the operation status. Obviously this is not very helpful, as the user would have to learn the instructions by heart or write down the help information on a piece of paper.

Present-day technology makes more intelligent on-line help systems possible, but such systems have usually only been applied in programs and products with a GUI and hardly in systems with a CLI. It is therefore interesting to try and find out whether it is possible to make CLI for small products with displays that are more efficient to work with, by offering information (descriptions of the functions) and instructions (actions to perform) on-line and by offering interactive on-line help.

10.3. Method

In this study, a menu-based interface, a command-line interface and a command-line interface with interactive help for a versatile telephone system with a four-line display are compared. The menu-based interface (guided system) has on-line help with only descriptive information, explaining what the various functions do; the procedural information is presented in the menus. The command-line interface has on-line help with both descriptive information and the operational instructions, telling what codes to enter.

[69] Various top managers from Microsoft have expressed their intention to focus on interactive help. See, for instance, the following quote from an interview with the present CEO Steve Ballmer (Computable, 31-3-2000, pp. 44-45): 'An innovation that Bill Gates is very interested in, is the Office Assistant. You type a question in ordinary language and the answer is prompted. But why would the assistant not simply do it? If that would be possible, what is then still the use of menus as they are now?'

All help in these telephones is offered in a purely textual form, including some symbols that were used on the telephone buttons (e.g., * and #). However, help can be offered in a number of ways for each type (guided or command-line). Help can be offered via a complete alphabetical list (index), thematic (table of contents, based on a hierarchical ordening of the various functions) or context-sensitive (help items are presented depending on the status of the machine). Here, the thematic presentation will not be further included, because it is difficult to combine the features of the telephone in appropriate and intuitive groups; some functions that can be grouped are automatically also organized alphabetically (call forwarding, call forwarding after four rings, call forwarding during absence).

Also, the help consists basically of two parts. One part explains what a certain item means. This help information is called 'descriptive', because it gives descriptions of the selected items. Another part of the help is 'procedural', because it presents the procedures that the user has to perform. Usually the help presents both descriptive information and procedural instructions. A question here is whether to present the description first and then the procedure or vice versa. It seems obvious to first explain and then instruct - and this is also common in many help systems of well-known software on computers. However, one could argue that the goal of users is to find the instruction, not the information. Especially if the user does have some idea of the function he wants to activate, but does not know how, it may be more appropriate to give the instruction first and then the explanation. Moreover, users may learn much better by doing: once they have entered the right codes and the machine has responded, they may understand what the function really means (an 'aha!' experience).

This analysis description results in eight versions of the help system: three versions with a guided system (menus), three versions with a command-line interface (enter codes), and one version with a command-line interface with interactive help.

1. A guided system with alphabetically organized help consisting of descriptive information followed by procedural instructions.
2. A guided system with alphabetically organized help consisting of procedural instructions followed by descriptive information.
3. A guided system with context-sensitive help (alphabetically ordered if there are various items to be presented. If various

possibilities have to be presented, the items will be ordered alphabetically).

4. A command-line interface with alphabetical help consisting of descriptive information followed by procedural instructions.
5. A command-line interface with alphabetical help consisting of procedural instructions followed by descriptive information.
6. A command-line interface with context-sensitive help; if various possibilities have to be presented, the items will be ordered alphabetically.

7. A command-line interface with interactive help; help is presented in the same way as in version 6 (context-sensitive, first description, then procedures), but an 'Execute' button is included.
8. A command-line interface with interactive help; help is presented in the same way as in version 6 (context-sensitive, first procedures, then description), but an 'Execute' button is included.

The telephones developed for this test are based on an existing (hardware) telephone that has a display but no on-line help. The display is only used for feedback. Programming the many functions has to be done by entering the codes given in the user manual. As a result, many users find it too difficult to program the functions and simply do not use them at all.

On the basis of this purely code-based telephone with a printed user manual, a completely new telephone with exactly the same features was developed on a computer screen. This telephone was developed with the eight versions of user interface and help systems that were described above: three versions with a guided system (menus) (figure 10.3), three versions with a command-line interface (figure 10.4) and two versions with a command-line interface with interactive help (with an 'Execute' button, figure 10.5) (p.193). The appearance of all test versions is identical; only interaction (guided via menus or applying a command-line interface) and the help systems differ. The versions with a guided system always present information on the display (menus from which the user can select); the versions with a command-line interface only present some feedback (e.g., what has been entered) and help information if the help button is pressed.

All versions of the telephone were programmed in Visual Basic and were fully operational on the computer screen by using the mouse, except that no connections could be made. They all had exactly the same features as the existing telephone and produced the same type of feedback and sounds. Some 'answers' that would be given after successfull completion of a task, were pre-programmed.

For each version five subjects were asked to perform nine tasks of increasing in difficulty (table 10.1).

Table 10.1. Tasks to be performed on the telephones.

1	Change the bell tone to melody 7 and speed 9.
2	Program the telephone so that the secretary at phone number 7179 can hear the ringing and answer the call after four rings.
3	Deactivate the function that you just programmed.
4	You are talking to someone at 4667; put this call through to 7179.
5	You receive a call and try to put it through to 7179, but this appears to be busy. Program your phone in such a way that 2999 will be called automatically once the line is free.
6	Program phone number 0015-42787179 under speed-dial button 12.
7	Program this telephone in such a way that no calls can be made (pin code = 123)
8	De-activate the function that you just programmed (pin code = 123)
9	Change the bell tone to melody 9 and speed 7.

The terminology of the tasks was in most cases different from the terms used in the feedback and help information on the display. For instance, the task was not formulated as 'Set Follow me if absent', but as 'Program the phone so that all calls will be diverted to another phone after four rings'. This is a realistic situation, as users in many cases formulate the problem themselves and then try out whether it is possible on the machine.

In total 40 subjects have participated in this test: 5 subjects for each version. Subjects were all students at the Delft University of Technology, an equal distribution of males and females, and mostly 3rd and 4th year students from the Faculty of Industrial Design Engineering. Each subject received ƒ 10,- after completing the test. For each subject half an hour was scheduled; a pretest had indicated that this would be sufficient and for the test this time scheme did not cause any problems.

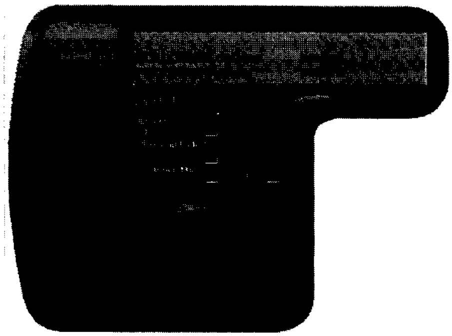


Figure 10.3. The menu based telephone set. Help is both descriptive and procedural. [Text in the 'display' has been translated into English]

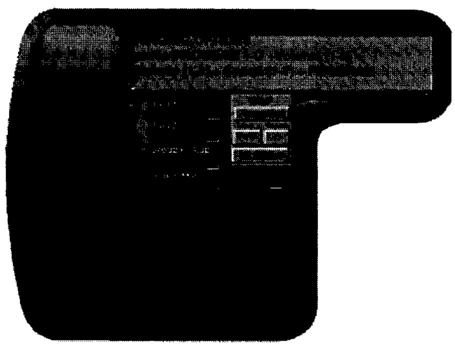


Figure 10.4. The telephone set with the CLI-interface with standard on-line help. The Help system consists of 'On' and 'Off', 'Previous' and 'Next' (<< and >>) buttons. Help is procedural and descriptive. [Text in the 'display' has been translated into English]

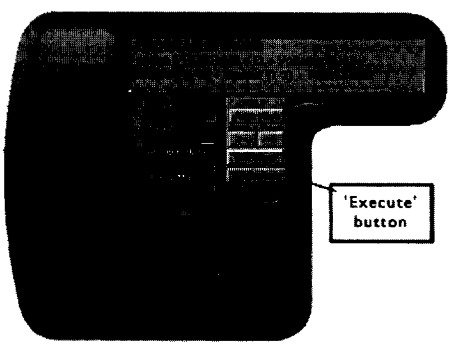


Figure 10.5. The telephone set with the CLI interface with interactive on-line help. The Help system consists of 'On' and 'Off', 'Previous' and 'Next' (<< and >>) buttons plus the 'Execute' button. [Text in the 'display' has been translated into English]

The computer measured the times that the subjects needed to perform the tasks successfully, the buttons that were pressed (= activated with the mouse) and the times between these activations. Calculations and graphical presentation were done using Microsoft Excel; a statistical analysis of the results was done using SPSS.

In this report only the relative efficiency of guided systems (the three menu versions combined) and the efficiency of command line systems (the three code versions combined) and the efficiency of applying interactive help in a help system for code-based systems will be discussed. Therefore, the results of the first three versions (15 subjects) are taken together and compared with the results of the 15 subjects who used the code-based on-line help systems and the 10 subjects who used interactive help for the code-based systems.

10.4. Results

In this test subjects needed on average 807 seconds (13-14 minutes) to perform the nine tasks successfully. An experienced user could perform these tasks in less than three minutes; the rest of the time was needed to find out how to do it (10-11 minutes). Subjects needed most of the extra time to find the right information. Applying it did not seem to take much time. Almost all subjects needed the help. Only a few of the subjects using the menu versions managed to (de)activate the functions without using help. The results of the tests are shown in table 10.2 and figure 10.6.

Table 10.2. Total times, means and standard deviations for three versions of the on-line help system on the telephones.

	guided	code (CLI)	code (CLI) plus interactive help
Mean	640	971	820
Standard deviation	310	285	169

A one-way ANOVA analysis shows that there is a significant difference between the three systems in the times required by the users to perform the tasks ($F = 4{,}9145$, $df = 2$, $p < .0138$). Comparisons between the individual eight versions did not present significant differences.

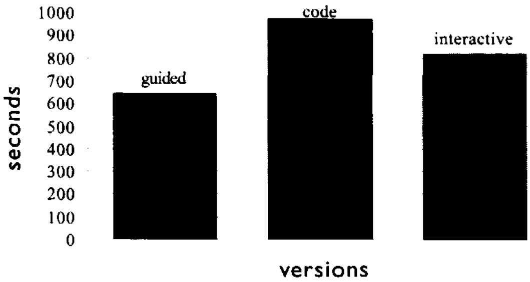


Figure 10.6. Mean times (in seconds) for execution of the tasks on the telephone sets with a guided (menu) system, a code system with standard help and a code system with interactive help.

10.5. Conclusions

The results show a significant difference between a guided system (menu-driven telephone interface) and a command-line interface (code-driven telephone interface) (on average 640 versus 971 seconds, a difference of 331 seconds). Guided systems are significantly faster than code-based help systems (t-test guided (n = 15) versus code (n = 15): $p < .005$, df = 28) for novice users (i.e., users who have to find codes they do not know by heart). This is, of course, exactly what we expected and is generally known (Preece, 1994).

Subjects using the guided (menu-driven) telephones also needed more time to complete the tasks than the command-line telephones with interactive help. But the difference is much smaller and statistically not significant: on average 820 versus 640 seconds, a difference of 180 seconds (t-test guided (n=15) versus interactive help (n = 10): $p < .249$, df = 18). So there is some indication, but no evidence, that interactive on-line help may be an improvement for systems with command-line interfaces.

Subjects needed less time to complete the tasks with the interactive help versions than with the standard help versions (820 versus 971 seconds, a difference of 151 seconds), but the difference is not significant (t-test: code (n = 15) versus code with interactive help (n = 10): $p < .297$, df = 18). Significant difference could appear if more subjects were to do the test.

This test also showed some problems with the interactive on-line help system that may have influenced the results. Some users in the beginning simply did not realize the function of the 'Execute' button and as a result did not use it, and they simply entered the codes manually, just like in the non-interactive versions. Once subjects discovered the possibilities, they performed the tasks very rapidly. Therefore, the overall impression is that interactive help can be very efficient, also for command-line interfaces, provided that the users understand the interactivity. It must be remembered here that all subjects are novice users of this system; once they get used to this new type of presentation, they may work faster using the interactivity. These tests were done in 1996, when interactive help was still very unknown.

Another problem in this test was that subjects sometimes tried to click on the 'display' of the telephone on the screen. Of course, this did not have any effect, because with a real physical phone one would not press with the fingers on the display to activate or de-activate a function. This confusion was probably caused by the idea that the phone used in this test was supposed to be an on-screen representation of a physical telephone and not an actual on-line telephone. This may also have influenced the test results, because this happened more often with subjects performing the tasks on the command-line telephones. Subjects using the guided systems were usually prompted to press numbers, which of course was done on the buttons.

For novice users, a guided system is the best solution, compared to command-line interfaces. Applying the procedural information is a secondary problematic element. Therefore, in products with a command line interface an interactive help may be an improvement, if users understand the usefulness of the function. Explaining this is a different matter. Maybe that is why all versions of WordPerfect for Macintosh since version 2 no longer had this function.

CHAPTER 11

Conclusions

The increasing complexity of society and the development of technological artefacts have evoked an increasing need for instructions and the development of media to represent these instructions. Instructions started in pictorial form; the pre-historic cave paintings in Spain and southern France may be examples, while the ancient Mesopotamian floor plan for a fortress is certainly an example. For thousands of years such pictures may have been the only media for instructing, apart from oral instructions, of course. The advent of abstract symbols for writing, starting about 10,000 BC, in the area of Sumeria (Bouwhuis, 2000, p. 102), also brought textual instructions. Since the invention of printing, many books have been produced with generic instructions, especially in military, technical and medical areas, but also on the art of painting, dancing, fishing, etc. In these books we can discover a wide variety of creative uses of pictures, both in their design and in the combinations with texts. Some instructional elements were also used in these books, especially reference systems with letters and numbers and reference lines. The Industrial Revolution brought large numbers of identical technical products, for which printed instructions were needed that gave specific instructions for each machine. These had to be reproduced in large numbers, very rapidly and at low cost. As a result, the early day manuals for sewing machines, typewriters, telephones and cars usually had rather few pictures, and these were mostly overview drawings of the products or major elements, separated from the texts. World War II has given a major push towards the use of more pictures and more instructive elements (arrows, dotted lines, prohibition signs, etc.). Also, the use of series of small pictures indicating the relevant details in procedures increased rapidly after World War II. Other major visual changes are the more general use of at least one supporting color in instructions and the influence of cartoons. A major change came when the possibilities of the computer could be applied for instructions. On screens the use of full color is common, and next to text and pictures, other media could be applied, such as animations and sounds.

New technologies may be available for instructing, and new media can be applied, but that does not necessarily mean that they are better.

Especially within the complex context of user instructions (as discussed in chapters 3 and 4) and the complex interaction and communication process (chapter 5), it is difficult to precisely define what media may be more efficient than others for user instructions. To be able to draw some generalizable conclusions about the relative efficiency of various media in user instructions, the Bieger and Glock taxonomy of specific types of information in user instructions was introduced in chapter 6 and tested in chapters 7,8 and 9).

11.1. Relative efficiency of text, pictures and animations

What does this study reveal about the relative efficiency of text, pictures or animations for operational, spatial, contextual and covariant information in instructions for products with a two-dimensional user interface?

Flowcharts (logical pictures) appeared to be a very inefficient medium for presenting operational and covariant information, compared to text and pictures; for contextual information the results indicate the same, but the differences with the other media were not significant (chapter 7). Given the fact that all subjects were students from a technical university with some experience in reading flowcharts, it is likely that this conclusion is certainly true for most users of modern technological products, since they probably have less experience. It is interesting to note that flowcharts are such an inefficient medium, because operational information is sometimes presented in user instructions, especially for telephones and some other telecom products. These are probably based on technical schemes from the electronic engineers who have designed the devices. One can, of course, question the conclusion, because other designs of the flowcharts used in this test may be more efficient. In this test, the flowcharts were designed according to the ISO standards - and these are not necessarily the best. Wright & Reid (1973) found flowcharts to be an efficient alternative to text in instructions for tasks that involved repetitive decision-making. But they did not motivate their design choices; maybe they intuitively found a better design.

Leaving out covariant and contextual information proved to be quite efficient in this test; therefore, it was decided to focus on the relative efficiency for operational and spatial information. In the test presented in chapter 7, pictures appeared to be more efficient for presenting operational information than text. This is in contradiction with the results from the tests discussed in chapters 8 and 9. In both these tests, text appeared to be the most efficient medium, pictures less efficient and animations the least efficient medium.

How does one explain this contradiction? Style differences could be an explanation, comparable to the variations in designs of the flowcharts that might explain variations in findings. Maybe the text in one of the tests could have been improved. Another difference is that the instructions in the first test (chapter 7) were presented on paper, whereas the instructions in the other tests (chapters 8 and 9) were all presented on-line. But there is a more important difference between these tests. In the first test (chapter 7) the text version contained all types of information (operational, contextual, covariant, etc.); the texts in the other two tests (chapters 8 and 9) presented only operational information and exactly this information was then compared with presentations in other media (pictures and animations). The latter two texts were also much shorter. In a comparison with all information present, the results of the first test are more relevant. In a comparison with only one type of information (operational) presented in different media, the conclusions from the latter two tests are more relevant.

Animations proved to be a relatively inefficient medium to transfer operational information. However, the differences with texts and pictures became smaller when the tasks had to be repeated. This effect was already visible when the nearly identical tasks had to be done at the end of the tests, less than half an hour later. It became all the more noticeable when the subjects had to repeat the tasks one week later: animations were still the least efficient, but the differences with text and pictures became small.

Consulting the instructions for a product takes time: one has to read and understand them before trying them on the product's user interface. Compared to trial-and-error learning, this requires an extra investment in time and cognitive load from the users, especially with the animated instructions. Subtracting the times needed just to consult the instructions from the total times needed to fulfill the tasks leads to totally different results for animations; here the animations appeared to be the most efficient. This may produce a guideline for the design of instructions: if the user will have to repeat the procedures often and if the user is prepared to invest extra time during their first use, it may be better to use animations. The primary time investment is high, but it will pay off if the tasks have to be repeated often, because the learning seems to go faster when animations are applied for operational information. A drawback is that the investment (consultation of animated user instructions is an extra difficult task) is required at the moment when the cognitive load is high because of the complexity of the product itself. Of course, it should be kept in mind that this conclusion is only valid for the operational information in instructions; nothing can be concluded for the other types of information or the combination.

Realistic (representational) pictures for operational information scored better than flowcharts and text in the first test (chapter 7), where all versions included all relevant types of information basically in text. In the second and third tests (chapters 8 and 9) pictorial presentation scored better than animations, but less good than textual presentation of this type of information. However, when the subjects had to perform the identical four tasks again within the same test (less than half an hour later), the results are different: pictures scored best, although there was no significant difference with textual presentation. One could conclude that the short-time learning curve is steeper with pictures. When the same subjects had to repeat the same tasks one week later, using the same instructions presented in the same medium, the effect was a little bit more noticeable, but not much: pictures appeared to be the most efficient medium for presenting operational information. In this test one week later, there was no significant effect within the test (identical tasks at the end of the test). The overall conclusion is that pictures are less efficient than text for presenting operational information for first use of the product. When learning is important pictures score better, although the improvement is not as significant as with animations.

Adding spatial information in separate pictures to help users locate the elements that they need, appeared not to be efficient for 2D products. The results were not significant in each of the separate tests, but both tests (7 and 8) produced the same results. Therefore, the results may give an indication but no proof. In these tests the spatial information was presented separately; nothing can be concluded about pictures that present, for instance, both operational and spatial information.

Can these conclusions be generalized? Only if the test subjects are representative for all users; if the designs of texts, pictures and animations are representative for all instructive texts, pictures and animations; if operational and spatial information can be the only types of information presented; and if the difficulty of the tasks is representative for tasks users have to perform on modern electronic consumer and office products, etc. At this moment there may be too many 'ifs' to generalize the conclusions. Further studies, with: other subjects, limited tasks and, therefore, better controllable instructions, and more variations with other types of information could lead to conclusions that are more generalizable. This attempt to draw general conclusions about the relative efficiency of various media for user instructions is based on the most detailed analysis of information in user instructions available; the study was limited to only a few types of information and applied to a situation with limited alternatives as far as the pictures are concerned. It is difficult to see how to have a more detailed analysis and still draw generalizable conclusions. Maybe it is not possible to draw generalizable conclusions - and that may be the most important conclusion.

11.2. Bieger & Glock taxonomy of information types

What does this imply for the taxonomy of Bieger & Glock, which has been the basis of this study?

Many studies have compared the relative efficiency of various media that can be applied in user instructions (such as text and pictures). However, the conclusions of these studies cannot be generalized and therefore have no scientific relevance. Bieger & Glock (among others) have indicated that conclusions can only be generalized if the types of information in instructions would be specified, classified and described in terms of their relevant characteristics. Only then, it might be possible to make comparisons and draw generalizable conclusions, Bieger & Glock assumed. This assumption and their taxonomy of types of information in user instructions were adopted in this study.

However, there was some severe criticism (§ 6.2.2) because Bieger & Glock did not refer to comparable classifications of others. The design of their tests was questionable at some points (think-aloud procedures and training of raters may have influenced their rating, excluding inventory information). The pictures that Bieger & Glock used did not always specifically and solely present the information that Bieger & Glock supposed they had, and the style of some pictures could easily be improved. Their classification was not always unambiguous (temporal information is implicitly included by the order of the sentences and the pictures). Bieger & Glock analyzed one type of information (contextual) on a different level than the other types (contextual information was analyzed at the paragraph and complete picture level, for the other types of information the texts were analyzed on a word or sentence level).

In spite of this criticism, the Bieger & Glock taxonomy was adopted and it proved to enumerate very well recognizable types of information, also for user instructions of products with a predominantly two-dimensional user interface. Students trained as raters unanimously classified types of information they found in user manuals according to the taxonomy, just as Bieger & Glock reported. But there was one big difference with the way Bieger & Glock had applied their taxonomy: in this test, the analysis of all types of information was executed on a paragraph and complete picture level, as Bieger & Glock had done for contextual information only. Analysis on a word level or on the level of details in a picture appeared too difficult and not result in overall agreement. However, it was decided that the taxonomy proved useful for analyzing information types in user instructions for products with a two-dimensional user interface at this more abstract level.

Application of the taxonomy to user instructions of products with a 2-D user interface, led to the conclusion that four types of information are ubiquitous and five types occur less often, just as Bieger & Glock had found. However, in contrast to Bieger & Glock, inventory information was not found very frequently, but covariant information did occur frequently. In conclusion: the four types of information occurring frequently in user instructions for products with a two-dimensional user interface are operational, spatial, covariant and contextual.

Frequency of occurrence does not necessarily mean that these types of information are also the most vital. The test discussed in the chapter 7 indicates that it is more efficient not to include contextual information. The data also suggest that it is not useful to include covariant information, but the differences are not significant.

The tests discussed in chapters 8 and 9 indicate that it is not very useful to include spatial information (in this case, the subtype indicating the location of objects), at least not in separate pictures added to other pictures indicating the operational information. Applying pictures that combine operational and spatial information could be very efficient. Thus, it could be efficient to include spatial information and to do this in pictures.

In conclusion, Bieger & Glock found the four types of information that occurred most frequently also to be the four essential types; the tests applying the taxonomy to user instructions for products with two-dimensional user interfaces produced different results. Four types were found frequently (two types different from the list that Bieger & Glock got for three-dimensional products), but only operational information proved to be essential. Covariant information may also be essential, yet the data indicate a different conclusion. Presenting contextual information appeared to be inefficient just like presenting spatial information in separate pictures.

The types of information in the taxonomy appeared to be very well recognizable, but only on the more abstract level of complete paragraphs and pictures, not on the detailed level of words and sentences. In these tests, pictures could be far less arbitrary than the ones Bieger & Glock used (§ 6.1), but then still some aspects were chosen somewhat arbitrarily. Moreover and more seriously: also in the tests discussed in chapters 8 and 9, the pictures do not solely and specifically include operational information. For instance, temporal information is included for the order of the pictures and the numbering (or the order of presentation in the animations) and some spatial information (not very exact, but approximate) is included.

11.3. What can be concluded for designers of user instructions?

The increasing complexity of technological products implies more responsibility for both user interface designers and user assistance designers. A close interaction between these two, starting at an early stage of the design of a product, is important for the creation of products that can really be used completely (including all features) in an efficient and pleasant way. Until this moment, user assistance designers are involved too late to design the user instructions and other parts of the user assistance. Involvement of user assistance designers is also necessary, because too often user interface designers still feel insufficiently responsible for the user assistance and do not realize the impossible tasks they require of user assistance designers. Moreover, if user interface designers do not feel responsible for at least the overall design of the user assistance, their products may have user interfaces and user assistances that do not combine very well, both in content and in design (for an example of a mismatch, see Westendorp & Mijksenaar, 2001, p. 4 or Westendorp 2001, p. 4). It would be an improvement if user interface designers realize that their user interfaces, magnificent as they may be, are not enough to help users use all features of the product. Convincing or persuading them to see it this way cannot only be the responsibility of schools or universities; product managers in charge of the development of modern technological products also have a responsibility here.

Managers in companies that produce modern technological products should also realize that the user assistance cannot and should not be treated apart from other elements of communication and interaction with the buyers of the products. User assistance should not be considered as just after-sales information. Too easily, a bad design of the user instructions may destroy the brand name that has been built up by investing so much money on pre-sales information, especially for advertising. Moreover, the money saved on user instructions may then have to be spent on the helpdesk. User assistance should be considered as a part of the complete communication and interaction process with those who bought the products. The user assistance could perhaps even be designed more in line with the commercial communication, or vice versa. Especially now that help can be offered on CD-ROMs and websites, the strict difference between pre-sales and after-sales information begins to fade. Potential buyers may have a look at the user assistance on the website before deciding. In addition, producers may refer the potential buyers to the user instructions on the website for an overview of the features and 'ease of use' may partially be demonstrated in a guided tour on the website, etc.

Some topics have been discussed that could result in a better integration of user interface and user assistance.
The first is that the user has to switch frequently between user

interface and user instructions. As a result of that, users may build up different mental representations of the product: a device model, if the user relies mainly on the graphical-spatial user interface (trial and error), and a procedure model, if the user relies mainly on the procedural-linear instructions. One could argue that a user interface based on the procedure model is sometimes preferable. For most products, such a user interface, based on procedures in user instructions is hard to imagine, but an interactive on-line help system that (partially) replaces the user interface (figure 3.4) is conceivable. Furthermore, recent developments in common software already show a further integration and more active role of the on-line help ('embedded help' or 'embedded assistance'). At the moment of this writing, such interactive help is still rather primitive. It may be a creative challenge to design a user interface in reverse order: taking the procedures, which present the operational information of the user instructions, as a starting point.

Secondly, it has been argued that users may make 'satisficing' choices ('this procedure may take me longer to get there, but I know it works') rather than optimizing choices ('what is the most efficient way to reach my goal?') when executing steps in a procedure (§ 5.4.2). For user interface designers and for user assistance designers this might in some cases lead to the decision not to include alternative routes to activate or de-activate features, or even not to design or present the quickest way possible. The consequence of this may be to design the user interface in such a way that users will have to follow procedures that are as identical as possible -omitting or hiding 'better' alternatives. 'The sooner users recognize a pattern for procedures the better' has to be the guideline.

Some more specific practical conclusions can be derived from the tests. The first is a starting point that Carroll formulated for the minimal manual: 'Less is more', the Bauhaus axiom, which has also been specified in this study. For user instructions for two-dimensional products it is efficient to present only the operational and perhaps the covariant information. Just instruct users what to do and possibly what the reaction to be expected from the device should be. It seems more efficient not to present contextual information - users get the idea of what the product or a specific feature can do for them by performing the procedures - or they may have that idea based on, for instance, commercial information. 'Learning-by-doing', Carroll's second starting point, may well be another useful guideline for user interface and user assistance designers. Of course, here 'learning-by-doing' means that all operational information should be presented, in contrast to Carroll's interpretation of the phrase.

The tests in this study also indicated that it is not necessary to include inventory information, at least it was not found very often. That is not very surprising, because for the construction of three-dimensional products it may be useful to know in advance what elements there are, or should be, and what tools to use. For programming telephones and using software this seems less relevant. More surprising may be the conclusion that users do not benefit from spatial information. However, this conclusion has to be limited to information that presents information about the location of elements in separate pictures. Covariant information was found abundantly in user instructions for two-dimensional products, and that is not surprising: the feedback from the machine is mentioned in the user instructions. It is surprising that the data from the test indicate that this does not help them very much, so that even this feedback information could be skipped in the user instructions. If this would be true (but the results are not significant), this means that users find the feedback from the device reliable enough and do not need much confirmation. It may be useful to bring to mind that the subjects in the test were all students with a technical background and with a lot of experience with, and interest in technical products. Little can be said about the five types of information that occur less frequent in user instructions.

For presenting just the operational information, text appeared to be the most efficient medium for immediate success. Pictures appeared less efficient, but the investment in time seems to pay off rapidly on repetitive usage of product and instructions (figures 8.11 and 9.1). For animations, this effect is much stronger: this medium proved to be very inefficient at first use, but more efficient with repetition. No specific advice concerning the choice of a medium can be given on the basis of these findings. Designers of user instructions who have to choose between these media must take into account the willingness of the user to accept an extra cognitive load and investment of extra time in order to get better results with repetition of the tasks. This is a complex advice, since it may vary for each product, for each feature and for each user. Moreover, it must be repeated that there are many 'ifs' in the conclusions of these tests.

A much easier advice to give user assistance designers is to look more in the past: the wheel of visual presentation of instructions has already been invented many times.

11.4. Mental processing of user instructions in various media

What do the argumentations (chapters 3, 4 and 5) and results (chapters 7, 8, 9 and 10) tell us about the user's perception and cognition when interacting with complex products while consulting user instructions? Only tentative conclusions can be drawn.

Text appeared to be more efficient for presenting operational information than pictures and animations, at least at first try. The most abstract presentation (text) is the fastest medium, and the most concrete presentation (animation) is the slowest. That is remarkable because in these tests (especially chapters 8 and 9) the pictures were exact representations of the part of the product that had to be used in every specific step. The picture in the user instructions is a direct copy of the part of the user interface to be used at each specific step. One would expect that recognition could not be easier and subjects would not have to make any mental conversion; cognitive load would be minimum in comparison to media that do require transformation. Why is this medium that allows direct recognition not more efficient than text, the medium that does require transformation?

One possible explanation could be that spatial information is missing in these tests, so that subjects had to search a little bit; integrated spatial information could have resulted in greater efficiency of pictures. On the other hand, adding spatial information did not lead to better results, and the conclusions remain the same if we compare the textual, pictorial and animated versions, all without spatial information. Thus, the results contradict the expectation that congruency between presentation of user instructions and user assistance would lead to the best performance, comparable to the results that Seel & Strittmatter (1984) had found (§ 5.5.A).

An explanation could be that the perception of just pictures or animations requires more time than perception of texts. This is obviously true for animations, and the results of a limited test (table 9.3) support this view. In this test, reading times for pictures were equal to those for texts. But for serious conclusions about the perception times of instructions, these calculated times are insufficient. This is also because this test does not really separate perception from further cognitive processing.

A more intriguing explanation would be that pictorial instructions of operational information are not cognitively processed in pictorial form, but have to be converted into another medium, for instance propositional form. This would contradict the 'mind's eye' hypothesis, which states that we process information in the same medium that we perceive it. It would support the view that information is processed in one medium, as Pylyshyn (1981) contended. The conclusion that instructions in the incongruent medium 'text' are more efficient than instructions in the congruent medium 'pictures' also supports the view that (instructional) information is not processed amodal (as Pylyshyn also put forward), but in a 'mode' that is closer to text (propositional). Within the context of these conclusions, the results of the test discussed in chapter 7 indicate that flowcharts (logical pictures) are definitely not the way we represent instructional information internally.

Pictorial instructions allow direct recognition, but have to be converted into another medium for a mental representation of the instruction; textual instructions do not allow direct recognition, but may be faster because this medium conforms more closely to our mental representation. If this conclusion is true, it supports the view presented in this study that users of modern technological products build up a procedure model, based on the procedures with instructional steps that they have to do to complete tasks, rather than a mental model or device based on the graphical-spatial user interface. Such a procedure model is based on repetition of procedures and is best viewed as an abstract pattern of series of steps (tables 5.2 and 5.3); it is difficult to imagine such an abstract pattern based on pictures. The abstraction of a picture or real object (e.g., a triangle) by itself is difficult to imagine, but the abstraction of various operational instructions is much more difficult to imagine.

The notion that user instructions, at least the operational information, are cognitively processed in one medium, independent of the medium in which they are presented to the user, supports the Soar theory, which also states that all information is processed in one, unitary way.

The Soar theory also states that learning is based on small, detailed pieces of information (cf., steps in an instructional procedure), plus a classification and recognition mechanism that allows abstraction. Here it is argued that such a system may be built up by the ongoing abstraction because of repetition of tasks in which a human being recognizes a pattern

The argumentations for processing instructional information in one medium and for learning by abstraction from repetition of procedures brings up the idea that cognition may, just like perception, be described in terms of pattern recognition and abstraction of patterns. That might implicate that cognition may be defined as just an (evolutionary) development, maybe following rules identical to perception. But this theorizing stretches far beyond facts and argumentations brought up in this study.

Aasman, J. (1995), *Modelling driver behaviour in Soar*. Doctoral thesis, Groningen University/KPN Research, Leidschendam.

Agricola, G. (1556), *De re metallica*. Basel. Woodcuts by H.R.M. Deutsch after Blasius Weffring. Translated in 1912 into English by the later president of the US, Herbert Hoover and his wife Lou Henry Hoover (paperback published in 1950).

Aichler, O. & Krampen, M. (1977), *Zeichensystemen der Visuellen Kommunikation*. Alexander Koch Verlagsanstalt, Stuttgart.

Anderson, J. (1983), *The architecture of cognition*. Harvard University Press, Cambridge.

Anderson, J. (1993), *Rules of mind*. Lawrence Erlbaum, Hillsdale.

Andre, E. & Rist, T. (1993), *The design of illustrated documents as a planning task*. In: Maybury, M.T. (ed.), Intelligent multimedia interfaces. AAAI Press/MIT Press, Menlo Park, pp. 94-116.

Anoshkey, A.M. & Catrambone, R. (1992), *Text and graphics in instructional design*. In: Tomek. I. (ed.), Computer assisted learning, Proceedings 4th ICCAL92 conference. Springer Verlag, Berlin, pp. 74-86.

Arcellana, J. (2000), *Designing assistance*. In: Westendorp, P., Jansen, C. & Punselie, R. (eds.), Interface design & document design. Rodopi, Amsterdam/Atlanta, pp. 51-66.

Arnheim, R. (1969), *Visual thinking*. University of California Press, Berkeley.

Arthur, P. & Passini, R. (1992), *Wayfinding: people, signs and architecture*. McGraw-Hill, London.

Bertin, J. (1981), *Graphics and graphic information processing*. Walter de Gruyter, Berlin.

Bertin, J. (2001), *Matrix theory of graphics*. Information design journal, 10(1), pp. 5-19.

Bieger, G.R. & Glock, M.D. (1984/5), *The information content of picture-text instructions*. Journal of experimental education, 53(2), pp. 68-76.

Bieger, G.R. & Glock, M.D. (1986), *Comprehending spatial and contextual information in picture-text instructions*. In: Journal of experimental education, 54(3), pp. 181-188.

Booher, H.R. (1975), *Relative comprehensibility of pictorial information and printed words in procedural instructions*. Human factors, 17(3), pp. 266-277.

Booker, P.J. (1963), *A history of engineering drawing*. Chatto & Windus, London.

Booker, P.J. (1964), *Engineering drawing practice*. In: Engineering heritage: highlights from the history of mechanical engineering. Vol. 1,

Heinemann, London, pp. 3-8.
Boorstin, D. J. (1991), *The discoverers*. Abrams, New York.
Bouwhuis, D.G. (2000), *Document-supported communication: design of new forms of interaction*. In: Westendorp, P., Jansen, C. & Punselie, R. (eds.), Interface design & document design. Rodopi, Amsterdam/Atlanta, pp. 101-119.
Bremer, A. (1998), *Technische product informatie*. [Technical product information] Course book IDE 1411. Delft University of Technology, Delft. [in Dutch]
Brockmann, R.J. (1996), *Victors W. Pagés early twentieth century automotive and aviation books*. In: Journal of business and technical communication, 10(3), pp. 285-305.
Brockmann, R.J. (1998), *From shipwrights to millwrights to the twenty-first century. Explorations in a history of technical communication in the United States*. Hampton Press, Creskill.
Bürdek, B.E. & Drukker, J.W. (1996), *Design: geschiedenis, theorie en praktijk van de produktontwikkeling*. [Design: history, theory and practice of product development]. Ten Hagen & Stam, Den Haag. [in Dutch]
Carroll, J.M., Mack, R.L., Lewis, C.H., Grischowsky, N.L. & Robertson, S.R. (1988), *Exploring a word processor*. In: Doheny-Farina, S. (ed.), Effective documentation; what we have learned from research. MIT Press, Cambridge, pp. 73-102.
Carroll, J.M. (1990), *The Nurnberg funnel: designing minimalist instructions for practical computer skills*. MIT Press, Cambridge.
Chase, W.G. & Clark, H.H. (1972), *Mental operations in the comparison of sentences and pictures*. In: Gregg, L.W. (ed.), Cognition in memory and learning. Academic press, New York.
Clark, R.E. (1983), *Reconsidering research on learning from media*. Review of educational research, 53(4), pp. 445-459.
Coe, M. (1996), *Human factors for technical communicators*. John Wiley & sons, New York.
Coppens, C. et.al. (1998), *En garde: schermen verbeeld*. [En garde: fencing in pictures]. Leuven University Press. Leuven. [in Dutch]
Corn, J.J., Horrigan, B. (1996), *Yesterday's tomorrows: past visions of the American future*. Johns Hopkins University Press, Baltimore.
Deacon, T. (1997), *The symbolic species*. Norton, New York.
DeKleer, J. & Brown, J.S. (1983), *Assumptions and ambiguities in mechanistic mental models*. In: Gentner, D. & Stevens, A.L. (eds.), Mental models. Lawrence Erlbaum, Hillsdale, pp. 155-190.
Dirken, J.M. (1999) (2nd edition), *Productergonomie* [Product ergonomics]. Delft University Press, Delft. [in Dutch]

Dixon, P. (1987a), *The structure of mental plans for following directions*. Journal of experimental psychology: learning, memory, and cognition, 13(1), pp. 18-26.

Dixon, P. (1987b), *The processing of organisational and component step information in written directions*. Journal of memory and language, 26(1), pp. 24-35.

Dixon, P. (1987c), *Actions and procedural directions*. In: Tomlin, R. (ed.), Coherence and grounding in discourse. John Benjamins, Amsterdam.

Duffy, T.M., Palmer, J.E. & Mehlenbacher, B. (1992), *On-line help: design and evaluation*. Ablex, Norwood.

Dwyer, F.M. (1972), *A guide for improving visualized instruction*. Pennsylvania State University Press, Learning service division, Pittsburgh.

Elders, F. (1971), *Arne Naess, biografie & filosofie* [Arne Naess, biography & philosophy]. International school of philosophy, Amersfoort. [in Dutch]

Feiner, S.K. & McKeown, K.R. (1990), *Generating coordinated multimedia explanations*. Proceedings of the 6th IEEE Conference on artificial intelligence applications, pp. 290-303.

Feiner, S.K. & McKeown, K.R. (1993), *Automating the generation of coordinated multimedia explanations*. In: Maybury, M. T. (ed.), Intelligent multimedia interfaces. AAAI Press/MIT Press, Menlo Park, pp. 117-139.

Floor, J.M.G. & Van Raaij, W.F. (1998), *Marketing-communicatie-strategie*. [Marketing communication strategy]. EPN, Houten. [in Dutch]

Frederiksen, C.M. (1975), *Representing logical and semantic structure of knowledge acquisition from discourse*. Cognitive psychology, 108(4), pp. 316-355.

Gagne, R.M. & Briggs, L.J. (1979), *Principles of instructional design*. Holt, Rhinehart & Winston, New York.

Gentner, D. & Stevens, A.L. (eds.), *Mental models*. Lawrence Erlbaum, Hillsdale.

Glass, A.L. & Holyoak, K.J. (1986) (2nd edition), *Cognition*. Random House, New York.

Goldsmith, E. (1984), *Research into illustration: an approach and a review*. Cambridge University Press, Cambridge.

Gombrich, E. (1990), *Pictorial instructions*. In: Barlow, H., Blakemore, C. & Weston-Smith, M. (eds.), Images and understanding. Cambridge University Press, Cambridge, pp. 26-45.

Gregory, R.L. (1995) (4th edition), *Eye and brain: the psychology of seeing*. Oxford University Press, Oxford.

Hering, H. (1997), *Historische Facetten der technischen Dokumentation*. In:

Kallnich, J. & Schwender, C., Erst lesen, dann einschalten! Zur Geschichte der Gebrauchsanleitung. Museum für Post und Kommunikation, Berlin, pp. 106-117.

Herrlinger, R. (1967) (2nd edition), *Geschichte der medizinischen Abbildung*. Vol. 1. Von der Antike bis um 1600. Moos, München.

Hoek, T. & Kaufman, L. (2000), *Help yourself: designing help to fit software users and their work*. In: Westendorp, P., Jansen, C. & Punselie, R. (eds.), Interface design & document design. Rodopi, Amsterdam/Atlanta, pp. 27-38.

Hoffman, D.D. (1998), *Visual intelligence: how we create what we see*. W.W. Norton & Company, New York & London.

Horton, W. (1991), *Illustrating computer documentation: the art of presenting information graphically on paper and on-line*. Wiley, New York.

ISO (1985), *Information processing. Documentation symbols, conventions for data, program network charts, system resources charts*. ISO standard 5807. International organization for standardization, Geneva.

Jakobson, R. & Halle, M. (1956), *Fundamentals of language*. Janua Linguarum 1, Mouton & Co. 's Gravenhage.

Jansen, C.J.M. & Steehouder, M.F. (1989), *Taalverkeersproblemen tussen overheid en burger*. [Linguistic communication problems between government and citizens]. Doctoral thesis, Utrecht University. SDU, The Hague. [in Dutch]

Johnson-Laird, P. (1983), *Mental models*. Harvard University Press, Cambridge.

Johnson-Laird, P. (1988), *How is meaning mentally represented?* In: Eco, U., Santambrogio, M. & Violi, P., Meanings and mental representations. Indiana University Press, Bloomington, pp. 99-119.

Kearsley, G. (1998), *The theory into practice database*. http://tip.psychology.org

Keren, G. (1996), *Technology, society and myopia*. Inaugural speech, Eindhoven University of Technology, Eindhoven.

Kieras, D.E. & Bovair, S. (1984), *The role of a mental model in learning to operate a device*. Cognitive science, 8(3), pp. 255-273.

Kist, B. (ed.)(1999), *The exercise of armes: all 117 Engravings from the classic 17th-century military manual by Jacob De Gheyn*. Dover, Dover Publications.

Knabe, K. (2000), *Designing on-line help for the Mac OS*. In: Westendorp, P., Jansen, C. & Punselie, R., Interface design & document design. Rodopi, Amsterdam/Atlanta, pp. 39-50.

Kotler, P. (1991) (7th edition), *Marketing management: analysis, planning, implementation and control*. Englewood Cliffs, New Jersey.

Kösler, B. (1990), *Gebrauchsanweisungen richtig und sicher gestalten*.

Forschungsergebnisse für die Gestaltung von Gebrauchsanweisungen. Forkel Verlag, Wiesbaden. [in German]

Krikke, J. (2000), *Axonometry: a matter of perspective*. IEEE Computer graphics and applications, July, pp. 7-11.

Lentz, L. (1998), *Tot nut van de Nederlandsche scheepsneringe. Een handboek voor de scheepsbouw uit de gouden eeuw*. [For use of the Dutch shipbuilding. A handbook of ship-building of the Golden Age]. Tekst[blad], 4. pp. 2-7. [in Dutch]

Levie, W.H. & Lentz, R. (1982), *Effects of text illustrations: a review of research*. Educational communication and technology journal, 30(3), pp. 195-232.

Lewis, C. & Rieman, J. (1994), *Task-centered user interface design: a practical introduction*.
ftp://ftp.cs.colorado.edu/pub/distribs/clewis/HCI Design-Book/

Lodges, M. (1994), *A treatise on fishing with an angle*. Journal of technical writing and communication, (24), pp. 37-48.

MacDonald-Ross, M. (1989), *Towards a graphic ecology*. In: Mandl, H. & Levin, J.R. (eds.), Knowledge acquisition from text and pictures. Elsevier science publishers, Amsterdam.

Maes, F., Ummelen, N. & Hoeken, H. (1996), *Instructieve teksten: analyse, ontwerp en evaluatie* [Instructive texts, analysis, design and evaluation]. Coutinho, Bussum. [in Dutch]

Mandl, H. & Levin, J.R. (1989 (eds.), *Knowledge acquisition from text and pictures*. Elsevier science publishers, Amsterdam.

Markman, A.B. (1999), *Knowledge representation*. Lawrence Erlbaum, Hillsdale.

Maybury, M.T. (ed.) (1993), *Intelligent multimedia interfaces*. AAAI Press/MIT Press, Menlo Park.

McPharlin, P. (1942), *Roman numerals, typographic leaves and pointing hands. Some notes on their origin, history and contemporary use*. Especially the chapter 'The pointing hand', pp. 47-65.

Meindl, J.D. (1987), *Chips for advanced computing*. In: Scientific American, special issue 'Trends in computing', Vol. 1, pp. 99-107.

Mijksenaar, P. (1991), *Ontplofte beelden*. [Exploded views]. Lecturis, Eindhoven. [in Dutch]

Mijksenaar, P. (1997), *Visual function*. 010 publishers, Rotterdam.

Mijksenaar, P. & Westendorp, P. (1999), *Open here: The art of instructional design*. Thames & Hudson, London.

Mirel, B., Feinberg, S. & Allmendinger, L. (1991), *Designing manuals for active learning styles*. Technical communication, 38(2), pp. 75-87.

Molitor, S., Ballstaedt, S.P. & Mandl, H. (1989), *Problems in knowledge* *acquisition from text and pictures*. In: Mandl, H. & Levin, J.R. (eds.), Knowledge acquisition from text and pictures. Elseviers science publishers, Amsterdam, pp. 3-35.

Naess, A. (1966), *Communication and argument: elements of applied semantics*. Allen & Unwin, London.

Neurath, O. (1936), *International picture language*. Kegan Paul, London.

Newell, A. (1990), *Unified theories of cognition*. Harvard University Press, Cambridge.

Newell, A. & Simon, H. (1972), *Human problem solving*. Prentice-Hall, Englewood Cliffs.

Nordbotte, J.C. & Crosby, M.E. (1999), *The effect of graphic style on data model interpretation*. Information system journal, 9(2), pp. 139-155.

Norman, D.A. (1983), *Some observations on mental models*. In: Gentner, D. & Stevens, A.L. (eds.), Mental models. Lawrence Erlbaum, Hillsdale, pp. 7-14.

Norman, D.A. (1988), *The psychology of everyday things*. Basic Books, New York.

Odlyzko, A. (1999), *The visible problems of the invisible computer: A skeptical look at information appliances*. In: First Monday, 9(4), University of Illinois, Chicago. http://firstmonday.org/issues/issue4_9/odlyzko/index.html

Paivio, A. (1983), *The empirical case for the dual coding*. In: Yuille, J.C. (ed.), Images, memory and understanding; in honor of Allan Paivio. Lawrence Erlbaum, Hillsdale.

Passini, R. (1984), *Wayfinding in architecture*. Van Nostrand Reinhold, London.

Peeck, J.W. (1987), *The role of illustrations in processing and remembering illustrated text*. In: Willows, D.M. & Houghton, H.A. (eds.), The psychology of illustration. Vol. 1: basic research. Springer Verlag, New York, pp. 115-151.

Peeck, J.W. (1989), *Trends in the delayed use of information from an illustrated text*. In: Mandl, H. & Levin, J.R. (eds.), Knowledge acquisition from text and pictures. Elseviers science publishers, Amsterdam, pp. 263 277.

Plaumann, S. (1997), *Der maister sol auch kunden schreiben vn lesen: Das Feuerwerkbuch von 1420*. In: Kallnich, J. & Schwender, C. (eds.), Erst lesen, dann anschalten. Zur Geschichte der Gebrauchsanweisung. Museum für Post und Kommunikation, Berlin, pp. 54-60.

Preece, J. (ed.) (1994), *Human-computer interaction*. Addison-Wesley, Harlow.

Putscher, M. (1972), *Geschischte der medizinischen Abbildung: von 1600 bis zur*

Gegenwart. Band II. 2e Auflage. Moos Verlag, München.

Pylyshyn, Z.W. (1981), *The imagery debate: analogue media versus tacit knowledge.* Psychological review, 88(1), pp. 16-45.

Rasmussen, J. (1986), *Information processing and human-machine interaction: an approach to cognitive engineering.* North-Holland publishing, Amsterdam.

Rasmussen J. (1991), *Mental models and the control of action.* In: Tauber, M.J. & Ackermann, D. (eds.), Mental models and human-computer interaction. Vol. 1. North-Holland, Amsterdam, pp. 41-69.

Rector, M. (ed.) (2000), *Medieval combat: a fifteenth-century illustrated manual of swordfighting and close-quarter combat by Hans Talhoffer.* Greenhill Books/Lionel Leventhal, London.

Robin, H. (1992), *The scientific image: from cave to computer.* Abrams, New York.

Ross, K. (1978), *Codex Mendoza: Aztec manuscript.* Miller graphics/Productions Liber, Fribourg.

Rouse, W.B. & Morris, N.M. (1986), *On looking into the black box: prospects and limits in the search for mental models.* Psychological bulletin (100), pp. 349-363.

Rutherford, A. & Wilson, J.R. (1991), *Models of mental models: an ergonomist - psychologist dialogue.* In: Tauber, M.J. & Ackermann, D. (eds.), Mental models and human-computer interaction 2. North-Holland, Amsterdam, pp. 39-58.

Scaife, M. & Rogers, Y. (1996), *External cognition: how do graphical representations work?* International Journal of human-computer interaction, 45(3), pp. 185-213.

Schank & Abelson (1977), *Scripts, plans, goals and understanding.* Lawrence Erlbaum, Hillsdale.

Schmalhofer, F. & Kühn, O. (1990), *The psychological process of constructing a mental model when learning by being told, from examples, and by exploration.* In: Tauber, M.J. & Ackermann, D. (eds.), Mental models and human computer interaction. Vol.1. North-Holland, Amsterdam, pp. 337 360.

Schriver, K. (1989), *Document design from 1980 to 1989: the challenges that remain.* Technical communication, 37(4), pp. 316-331.

Schriver, K. (1997), *Dynamics in document design.* John Wiley & sons, New York.

Schumacher, G.M. & Waller, R. (1985), *Testing design alternatives: a comprehension of procedures.* In: Duffy, T. & Waller, R. (eds.), Designing usable texts. Academic Press, Orlando, pp. 377-403.

Schwender, C. (1997), *Erst lesen, dan einschalten: zur Geschichte der Gebrauchsanweisung.* Museum für Post und Kommunikation, Berlin.

Pp. 106-117.

Seel, N.M. & Strittmatter, P. (1984), *Stratiegien zum Erwerb geographischen Wissens und bildhafter räumlicher Vorstellungen.* Unterrichtswissenschaft 12 (1), pp. 33-47.

Seel, N.M. & Strittmatter, P. (1989), *Presentation of information by media and its effect on mental models.* In: Mandl, H. & Levin, J.R. (eds.), Knowledge acquisition from text and pictures. Elsevier science publishers, Amsterdam, pp. 37-58.

Shneiderman, B. (1998) (3rd edition), *Designing the user interface: strategies for effective human-computer interaction.* Addison-Wesley, Amsterdam.

Simon, H. (1955), *A behavioral modal of rational choice.* Quarterly journal of economics, 69(2), pp. 99-118.

Sims-Knight, J.E. (1992). *To picture or not to picture; how to decide?* Visible language 26(3/4), pp. 325-388.

Snoek, H. & Hekkert, P. (1998), *Directing designers towards innovative solutions.* In: Jerrard, B., Newport, R. & Trueman, M. (eds.), Managing new product innovation. Taylor & Francis, London, pp. 167-180.

Stone, D.E. (1980), *Reading text with pictures.* Paper presented at the annual meeting of the American Educational Research Association. Boston.

Stone, D.E. & Glock, M.D. (1981), *How do young adults read directions with and without pictures?* Journal of educational psychology, 73(4), pp. 419 426.

Stone, D.E., Pine, C.K., Bieger, G.R. & Glock, M.D. (1981), *Methodological issues concerning research on reading text with illustrations.* Department of Education, College of agriculture and life sciences, Technical report 2, Cornell University, Ithaca.

Strothotte, C. & Strothotte, T. (1997), *Seeing between the pixels: pictures in interactive systems.* Springer, Berlin.

Sweller, J. & Chandler, P. (1994), *Why some material is difficult to learn.* Cognition and instruction, 12(3), pp. 185-233.

Tabachneck-Schijf, H.J.M. & Simon, H.A. (1996), *Alternative representations of instructional material.* In: Peterson, D. (ed.), Forms of representation: an interdisciplinary theme for cognitive science. Intellect books, Exeter, pp. 28-47.

Tauber, M.J. & Ackermann, D. (1990), *Mental models and human-computer interaction 2.* North-Holland, Amsterdam.

Thölke, J. (1998), *Product feature management.* Doctoral thesis, Delft University of Technology, Delft.

Tolman, E.C. (1932), *Purposive behavior in animals and men.* Appleton Century-Crofts, New York.

Ummelen, M. (1997), *Procedural and declarative information in software manuals.*

Doctoral thesis. Twente University, Enschede.
Urcuioli, P.J. (2001), *Categorization & acquired equivalence.* In: Cook, R.G. (ed.), Avian visual cognition. www.pigeon.psy.tufts.edu/avc/toc.htm
Van der Meij, H. (1998), *Optimizing the joint handling of manual and screen.* In: Carroll, J.M. (ed.), Minimalism beyond the Nurnberg Funnel. MIT Press, Cambridge, pp. 275-309.
Van der Waarde, K. (1993), *An investigation into the suitability of the graphic presentation of patient package inserts.* Doctoral thesis. Department of typography & graphic communication, University of Reading, Reading.
Van der Waarde, K. (1999), *The graphic presentation of patient package inserts.* In: Zwaga, H., Boersema, T. & Hoonhout, H. (eds.), Visual information for everyday use. Taylor & Francis, London, pp. 75-81.
Van der Waarde, K. & Westendorp, P. (2002), *Design concepts of instruction manuals.* In: Schneider, A. (ed.), Information design. Graphic-Sha, Tokio. [In Japanese; to be published in English by John Benjamins, Amsterdam in 2003]
Van Hees, M. (1994), *Oud geleerd, oud gedaan. Optimalisering van handleidingen voor ouderen.* [Learned at old age, done at old age. Optimalization of user manuals for the elderly]. Graduation thesis, Utrecht University, Utrecht. [in Dutch]
Van Hees, M.M.W. (1996), *User instructions for the elderly: what literature tells us.* Journal of technical writing and communication 26(4), pp. 521-536.
Van Mourik, F. & Mijksenaar, P. (1998), *Pictogrammen: toepassing en ontwerp.* In: Ruijsch, P. & Theys, E. (eds.), Pictogrammen & iconen. [Z]oo/Vd Publishers, Eindhoven, pp. 4-15.
Van Mourik, F. & Westendorp, P.H. (2000), *The aspect of style in instructional illustrations.* Paper presented at the Expert Forum Manual Design, Mälardalen University, Eskilstuna, Sweden. www.visuelltforum.com/2000/utland/index.html.
Wagenaar, W.A. & Embrey, D.E. (1990), *Human error.* SIGTTO information paper 8, London.
Westendorp, P. (1993), *Gebruiksaanwijzingen.* [User manuals]. IVO, Eindhoven. [in Dutch]
Westendorp, P.H. (1994), *Design concepts of user manuals.* In: Steehouder, M., Jansen, C., Van der Poort, P. & Verheijen, R. (eds.), Quality of technical documentation. Utrecht Series in Language and Communication. Rodopi, Amsterdam, pp. 39-48.
Westendorp, P. et al. (1994), *Tekst, beeld of stroomdiagram voor technische*

instructies? [Text, pictures or flowcharts for technical instructions?]. Internal report VVI 1. Delft University of Technology. Delft. [in Dutch]

Westendorp, P.H. (1995a), *The subminimal manual?* In: Steehouder, M.F. (ed.), Disappearing borders. Forum95 proceedings. Intecom, Stuttgart, pp. 1-3.

Westendorp, P.H. (1995b), *From minimal manual to subminimal manual.* In: Steehouder, M.F. (ed.), Forum95 postharvest proceedings. Intecom, Stuttgart, pp. 1-5.

Westendorp, P.H. (1996a), *Text, pictures or animations in instructions for use: a validation of different media for specific types of information.* In: Ensink, T. & Sauer, C. (eds.), Researching technical documents. Groningen series on language and communication. Groningen University Press, Groningen. pp. 181-197.

Westendorp, P.H. (1996b), *Learning efficiency with text, pictures and animation in on-line help.* Journal of technical writing and communication, 26(4), pp. 401- 417.

Westendorp, P.H. (1996c), *Tekst, beeld of animatie voor on-line help bij een telefoonsysteem.* [Text, pictures or animation for on-line help in a telephone system]. Internal report VVI 2, Delft University of Technology, Delft. [in Dutch]

Westendorp, P.H. (1997a), *Guided systems versus interactive on-line help.* In: Morris, A., Houser, R. & DeLoach, S. (eds.), IPCC97 Proceedings. IEEE, Piscataway, pp. 1-8.

Westendorp, P.H. (1997b), *Leren van tekst, beeld of animatie voor on-line help bij een telefoonsysteem.* [Learning from text, pictures or animation for a telephone system]. Internal report VVI 3. Delft University of Technology, Delft. [in Dutch]

Westendorp, P. (1998), *Product design, interface design & help design.* Course book IDE 323, Faculty of Industrial Design Engineering. Delft University of Technology, Delft. [in Dutch] www.io.tudelft.nl/education/ide323

Westendorp, P.H. (1999a), *Text, pictures or flowcharts for specific types of information.* In: Zwaga, H., Boersema, T. & Hoonhout, H. (eds.), Visual information for everyday use. Taylor & Francis, London, pp. 83-89.

Westendorp, P. (1999b), *Design concepts of user manuals.* Workshop presented at the STIC meeting (Society of technical communicators), April 27, Design Academy Eindhoven. www.stic.nl

Westendorp, P. (2000), *Reality, theory and design aspects of instructions for use.*

Keynote address, Proceedings Expert Forum Manual Design, Eskilstuna, Sweden. www.visuelltforum.com/2000/utland/index.html

Westendorp, P. (2001), Design of visual instructions. Wednesday Proceedings of the ninth annual WinWriters On-line Help conference. Santa Clara, California, March 4-9. Pp. C84 1-13.

Westendorp, P., Jansen, C. & Punselie, R. (2000), *Parallels and dichotomies in interface design & document design*. In: Westendorp, P., Jansen, C. & Punselie, R. (eds.), Interface design & document design. Rodopi, Amsterdam/Atlanta, pp. 5-25.

Westendorp, P. & Mijksenaar, P. (2001), Creatieve instructies. [Creative instructions]. Product 9(5), pp. 4-9. [in Dutch]

Westendorp, P. & Van der Waarde, K. (2001), *Icons: support or substitute?* Editorial introduction. Information design journal 10(2), pp. 91-94.

Wickens, C.D., Sandry, D. & Vidulich, M. (1983), *Compatibility and resource competition between modalities of input, central processing and output*. Human Factors 25(2), pp. 227-248.

Williams, T.R. (1993), *Text or graphics: an information processing perspective on choosing the more effective medium*. Journal of technical writing and communication 23(1), pp. 33-52.

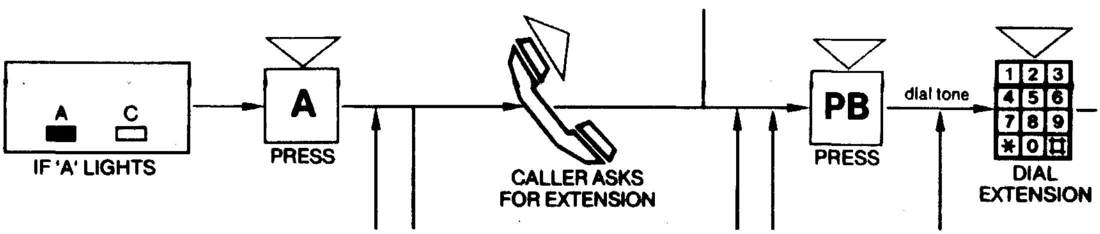
Winn, W. (1989), *Design and use of instructional graphics*. In: Mandl, H. & Levin, J.R. (eds.), Knowledge acquisition from text and pictures. Elseviers science publishers, Amsterdam, pp. 125-139.

Winn, W. & Snyder, D. (1996), *Cognitive perspectives in psychology*. In: Jonassen, D.H. (ed.), Handbook of research for educational communications in technology. McMillan, New York, pp. 112-142.

Wright, P. (1988), *Issues of content and presentation in document design*. In: Helander, M. (ed.), Handbook of human computer interaction. Elsevier science publishers, Amsterdam, pp. 629-652.

Wright, P. & Reid, F. (1973), *Some alternatives to prose for expressing the outcome of complex contingencies*. Journal of applied psychology, 57(2), pp. 160-166.

Young, R.M. (1983), *Surrogates and mappings: Two kinds of conceptual models for interactive devices*. In: Gentner, D. & Stevens, A.L. (eds.), Mental models. Lawrence Erlbaum, Hillsdale, pp. 35-52.



A
C
IF 'A' LIGHTS
A
PRESS
CALLER ASKS
FOR EXTENSION
PB
PRESS
dial tone
1 2 3
4 5 6
7 8 9
* 0 #
DIAL
EXTENSION

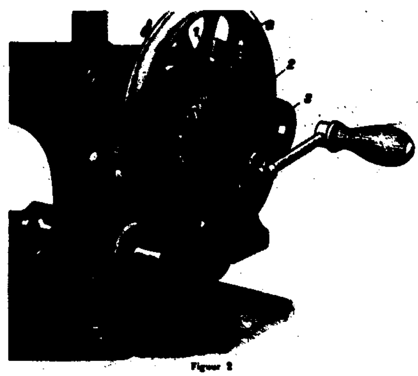
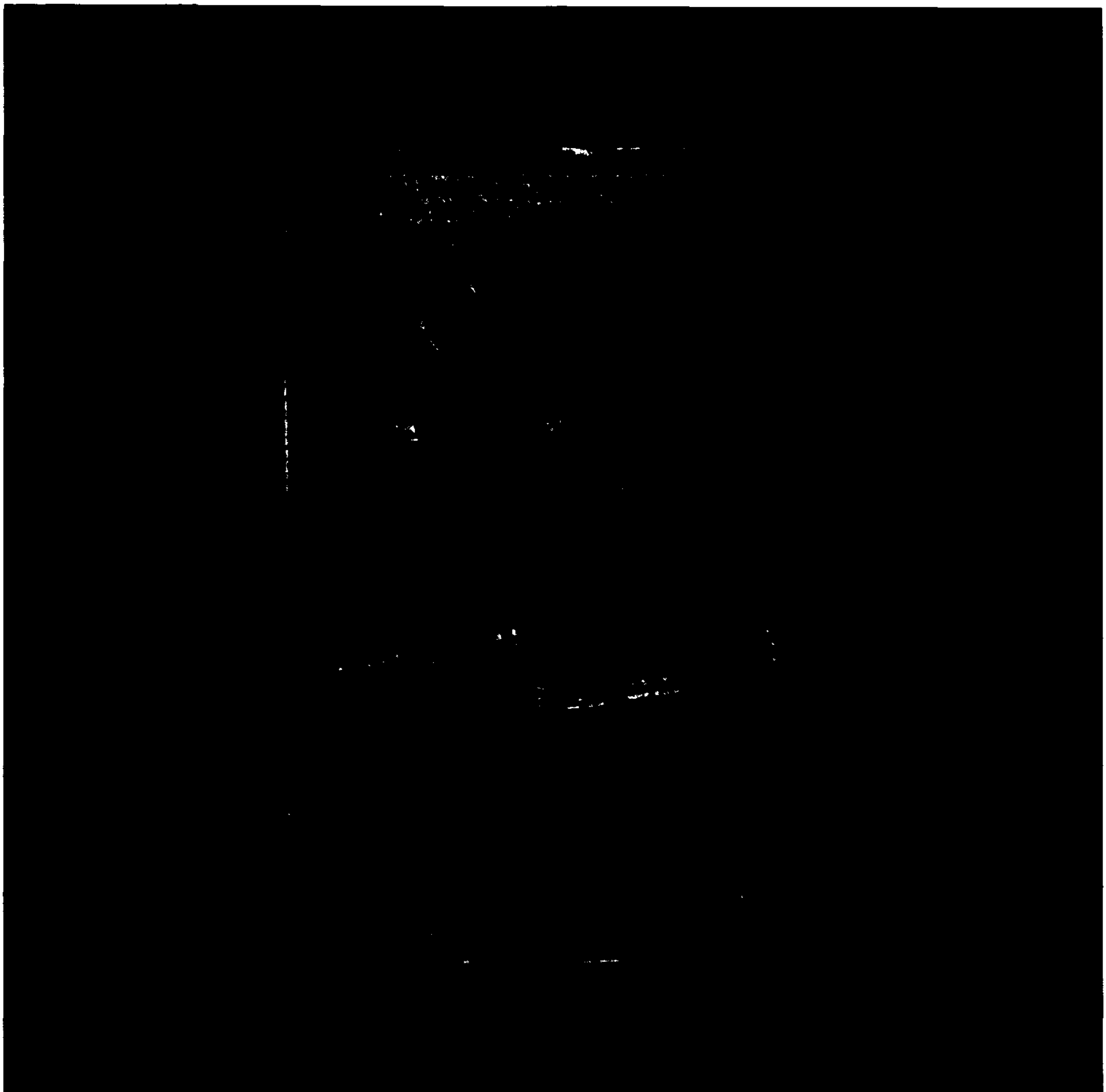


Figure 2

Presentation media for product interaction

Piet Westendorp



Presentation media for product interaction

[addendum]

Proefschrift
ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus, prof. dr. ir. J.T. Fokkema,
voorzitter van het College van Promoties,
in het openbaar te verdedigen op
maandag 10 juni 2002 om 16:00 uur
door

Pieter Hendrikus WESTENDORP
doctorandus in de Nederlandse Taal- en Letterkunde
geboren te Baarn

Dit proefschrift is goedgekeurd door de promotoren:
Prof. P.P. Mijksenaar
Prof. dr. J.M. Dirken Eur Ing.

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. P.P. Mijksenaar, Technische Universiteit Delft, promotor
Prof. dr. J.M. Dirken Eur Ing., Technische Universiteit Delft, promotor
Prof. dr. J. Aasman, Technische Universiteit Delft
Prof. dr. D.G. Bouwhuis, Technische Universiteit Eindhoven
Prof. dr. J.W. Drukker, Technische Universiteit Delft
Prof. dr. C. Strothotte, Hochschule Magdeburg, Duitsland
Dr. K.M. van der Waarde, Technische Universiteit Delft, adviseur

ISBN 90-9015900-2

Cover illustrations

Cataract operation. 12th Century. Manuscript Sloane. London Library (Herrlinger, 1967, p. 17).

Do not use lighters. US Air - Boeing 757 Safety card

Arrow on sewing machine. Naumann KL-14, late 19th century.

Philips TBX telephone operator manual. Flowchart.

Presentation media for product interaction

Summary

This study focuses on the relative efficiency of text, pictures and animation for some specific types of information in user instructions. Chapter 1 presents a general introduction with a motivation from theory and from practice for this study. Chapter 2 presents an overview of the increasing necessity, historic development and design aspects of user instructions. Here it is argued that technical (consumer) products are provided with more and more features, and are therefore rapidly becoming more complex and difficult to use. As a result, user instructions are becoming more important, in spite of all efforts to improve user interfaces.

The overview in chapter 2 shows that specific user instructions (for a specific type and model of a product) originated around the middle of the 19th century, when mass production of complex technical consumer products started. Both conceptual designs and detailing of the user instructions develop on the basis of generic instruction books that have been produced since the Middle Ages, especially for medical, technical and military procedures. New developments in visual instructions are discussed, including instructive elements (such as arrows, indicating hands and color for indicating parts). Conclusions for practical application about the relative efficiency of text, pictures and animation for user instructions must be judged within this context of possible concepts and detailing.

In the next three chapters (3, 4 and 5) the communication and interaction process between user and technical product is discussed; first on a sociological level (relation between producers and users of instructions), then on a meso-level (communication between the user of a product, the product and the user instructions), and finally on a psychological level (perception and cognition). The starting point here is that the user communicates and interacts with the products' user interface and with the supporting user assistance as well.

Chapter 3 describes the complex communication process between designers and users of user instructions. This chapter shows the overwhelming number of factors involved in each part of the process: making, transmitting and consulting the message. User instructions are not made by individuals, but by teams of professionals with a wide variety of backgrounds, such as lawyers, technicians, programmers, graphic designers, writers and illustrators. The target group may consist of hundreds of millions of people with a very wide variety of backgrounds, educations, technical inclination, etc. User assistance may consist of many different elements, such as advertisements, user instructions, on-line help systems and guided tours. Often the user will have made some kind of mental representation of a product before he bought it. Moreover, the 'message' in this process is often very complex: especially electronic products may have very many features. Possible practical conclusions of studies contrasting the relative efficiency of various media for instruction must also take these kinds of factors into account.

Chapter 4 zooms in on the communication and interaction process between user, product and user instructions. Too often the focus has been on just the relation between user and product (user interface). First the focus is on how often the users have to consult both the user interface and the user instructions for every detailed step in each procedure of the user instructions. Next, a model is developed which describes this switching between user interface and user instructions. The goal of this model is to map all user tasks in this triangle between the user and user instructions and the product.

This switching between user interface and user instructions is the starting point for a rather extensive discussion of the perception and cognitive processing of the interaction and communication of someone who tries to operate a complex technical product while consulting the user instructions (chapter 5). It is argued that perception and cognition processing of the graphical-spatial user interface, on the one hand, and the procedural-linear user instructions, on the other hand, could be quite different. It is further argued that users may develop mental representations that are rather more procedural-linear (analogue to the user instructions) than graphical-spatial (analogue to the user interface). If so, the mental representation is better depicted as an ongoing abstraction of procedures that users have to follow to activate or de-activate features of the product than a depiction based on the graphical-spatial user interface. Users have to recognize a linear pattern, rather than a graphical pattern, which is why it might be useful to try and design user instructions and user interfaces as congruent as possible. As a result of this reasoning, it is argued that user interface should be designed so that users can follow the same procedures for activating and de-activating features as much as possible.

Chapter 6 discusses the types of information that may occur in user instructions. From the argumentation by Bieger & Glock it follows that conclusions from tests contrasting the relative efficiency of various media for instructions can only be generalized if they are based on specified types of information in user instructions. The nine types of information that Bieger & Glock have discerned in assembling instructions are criticized, but yet they are used as a basis for tests. Then a discussion follows about the possibility to apply the Bieger & Glock taxonomy to user instructions for products with two-dimensional or nearly two-dimensional user interfaces, such as modern electronic products or software.

The next three chapters (7, 8 and 9) report about tests contrasting various presentation media for some specified types of information. In chapter 7, first the types of information that are dominant in user instructions for products with a (nearly) two-dimensional user interface are discussed. Where Bieger & Glock found inventory, operational, spatial and contextual information to be dominant (for assembly instructions), in this study operational, spatial, covariant and contextual information appeared to be dominant. A test with some variations of the presentation media for operational, contextual and covariant information in a complete manual for a telephone system indicated that covariant information (in these products usually feedback information) did not contribute to the efficiency. Contextual information (usually indicating what should be the result of a series of actions in a procedure) also seemed counter-productive rather than useful. Replacing text for the presentation of operational information with representational pictures led to a relatively efficient completion of tasks by subjects in the test. Replacing texts with flowcharts appeared not to be efficient for operational, nor for contextual nor for covariant information.

Chapter 8 presents a test in which either only operational information is presented to subjects who had to perform tasks on a telephone, or operational and spatial information. The operational information was presented in either text, pictures or animations; the spatial information only in separate pictures. The results indicate that text is the most efficient medium and animation by far the least efficient medium, both with or without the added spatial information in separate pictures. Presenting spatial information in separate pictures appeared to be very inefficient - subjects performed better with the versions without this information. The results also indicated that contextual information appears not to be required for a good understanding of the system; subjects performed very well without this information. A comparison of the first four tasks that subjects had to do and the nearly identical last four tasks indicated that the differences between the relative efficiency of text, pictures and animations had become much smaller within the same test that lasted at most half an hour. It seems that subjects had learned faster with the picture and animation versions.

The next chapter (9) describes a test that is identical to the one discussed in chapter 8, but in this test the same subjects were asked to repeat the same tasks using the same instructions in the same medium one week later. This test confirmed the conclusions of the previous test. Repetition of the tasks one week later showed that animation indeed proved to be more efficient on repetitive use of the product. Animation for operational information requires a big investment of time (compared to text and pictures), but it pays off with repetitive application of the same tasks. The time subjects needed for completion of the tasks decreased more than for subjects using instructions in other media. If the time needed just for consultation of the instructions were not taken into account, animations would be the most efficient medium when the tasks would have to be repeated. However, repetition of the nearly identical tasks at the end of the second test did not show any further improvement.

A totally different study is presented in chapter 10. Here it is argued that interactive on-line help could be beneficial for users of products with displays and screens. 'Interactivity' here means that the on-line help system does not just offer instructions, but also allows the user to immediately perform the instructions from within the help system. Thus, the on-line help would partially replace the 'regular' user interface. This study focused on products with small displays, where guidance via menus may not always be useful; products and software with screens were not included. It was presupposed that entering short codes would be more efficient if the user knew these codes by heart. A contrastive test with variations of a help system for a virtual telephone (with all functions working realistically) indicated that an interactive help system could indeed lead to a more efficient performance of the tasks for products when codes had to be entered in a command-line interface, but the difference was not significant. However, a menu-driven system appeared more efficient, at least for first time users.

The general conclusions are presented in chapter 11, with some remarks concerning the generalizability. It is further concluded that the Bieger & Glock taxonomy, which was the basis for the tests in chapters 7, 8 and 9, on the one hand produced very well recognizable types of information in user instructions, but on the other hand was also subject to severe criticism. Although the pictures used for instructions for products with a two-dimensional user interface can be far less arbitrarily chosen than pictures for assembly instructions, it was not always possible to fully separate the several types of information and offer just one type pictorially.

Next, some suggestions for practice are discussed. Especially on the basis of chapter 3, a plea is made for a more intense collaboration between user interface designers and user assistance designers. A further integration between user instructions and pre-sales information (such as advertising) is also made. On the basis of the argumentation in chapter 5, it is advised to design user interfaces to correspond as much as possible with the procedure model that users may develop during installation and use of the product. The difference between the optimizing choice and the 'satisficing' choice leads to the advice to accept that users are often content with a less efficient way to operate the product, if it is more standard, and to possibly design user interface and user instructions on this basis. Then some specific advice is presented concerning minimalization of manuals.

Finally an attempt is made to explain why the incongruent medium 'text' is more efficient during first use than the more congruent medium 'pictures'. It is suggested that pictures (for operational information in user instructions) are cognitively not processed as pictures, but have to be converted to a more abstract representation before users can execute the instruction. This would imply that information is not processed in a dual coded way, and that the 'mind's eye hypothesis cannot be confirmed. If this is the case, it would support the Soar theory, which assumes that humans process all information in one unitary way.

Presentatiemedia voor interactie met producten

Samenvatting [in Dutch]

Dit boek behandelt de relatieve efficiëntie van tekst, beeld en animatie voor enkele specifieke soorten informatie die voorkomen in gebruiksinstructies. Daartoe wordt - na de algemene introductie in hoofdstuk 1- eerst in hoofdstuk 2 een overzicht gegeven van de noodzaak, de historische ontwikkeling en de ontwerpaspecten van gebruiksinstructies. Beargumenteerd wordt dat de technische (consumenten)producten snel complexer worden omdat ze steeds meer mogelijkheden krijgen. Daarom worden gebruiksinstructies steeds belangrijker, ondanks pogingen om user interfaces van producten gebruiksvriendelijker te maken. Het overzicht laat zien dat specifieke gebruiksinstructies voor een speciaal merk en type product rond het midden van de negentiende eeuw ontstonden met de komst van seriematig geproduceerde technische apparaten voor een groot publiek. Zowel concepten als detailleringen hiervan vloeien voort uit de generieke instructies die sinds de middeleeuwen zijn gemaakt, vooral voor medische, technische en militaire handelingen. De nieuwe ontwikkelingen op het gebied van de visuele instructies worden behandeld, waaronder de instructieve elementen (zoals pijlen, handen die aanwijzen en gebruik van kleur om een deel aan te duiden). Eventuele op de praktijk gerichte conclusies ten aanzien van de relatieve efficiëntie van de verschillende media (tekst, beeld en animatie) voor gebruiksinstructies dienen gezien te worden binnen deze context van mogelijke concepten en detaillering.

Vervolgens wordt in een drietal hoofdstukken (3, 4, 5) het communicatie- en interactieproces tussen gebruiker en technisch product behandeld; eerst op sociologisch niveau (communicatie tussen makers en gebruikers van instructies), dan op meso-niveau (interactie van gebruiker met product en gebruiksinstructie) en vervolgens op psychologisch niveau ('het denken van den gebruiker'). Hierbij dient als uitgangspunt dat de gebruiker enerzijds communiceert en interacteert met de user interface van het product en anderzijds met de

user assistance, de aanvullende ondersteuning. Eerst (hoofdstuk 3) wordt het communicatieproces tussen producent en gebruiker van instructies beschreven. Dit hoofdstuk toont de hoeveelheid factoren die een rol spelen bij het maken, communiceren en consulteren van de gebruiksinstructies. De 'maker' is een team dat kan bestaan uit juristen, technici, programmeurs, schrijvers, tekenaars en vele anderen. De doelgroep kan uit honderden miljoenen mensen te bestaan, van velerlei achtergrond, opleiding, belangstelling, technische interesse, enzovoort. De assistentie voor de gebruiker kan bestaan uit een groot aantal middelen, zoals advertenties, de gebruiksaanwijzing, het on-line help-systeem en de 'rondleiding' (waarbij de meest interessante kenmerken van het product getoond worden). Bovendien is de 'boodschap' in dit proces vaak uitermate complex: met name elektronische producten en programmatuur kunnen honderden aanvullende functies te hebben. Eventuele praktische conclusies van onderzoek naar de relatieve efficiëntie van tekst, beeld en animatie voor instructie dienen ook rekening te houden met al deze factoren.

In hoofdstuk 4 wordt verder 'ingezoomd' op het communicatie- en interactieproces tussen gebruiker, product en gebruiksinstructies. Vaak wordt bij dit proces alleen acht geslagen op de gebruiker en het product. Hier wordt eerst beschreven hoe gebruikers van technisch complexe producten bij elke detailstap in een instructieprocedure zowel user interface als gebruiksinstructie moeten consulteren. Vervolgens wordt een model ontwikkeld dat dit 'heen-en-weer schakelen' tussen de user interface van het product en de gebruiksinstructie probeert te beschrijven. Het model heeft als doel alle taken van de gebruiker in deze driehoeksverhouding in hun onderling verband onder te brengen.

Dit 'heen-en-weer schakelen' tussen user interface en gebruiksinstructies vormt het uitgangspunt voor een beschouwing van perceptie en cognitieve verwerking van de interactie en communicatie van iemand die een technisch complex product probeert te gebruiken en daarbij de gebruiksinstructies consulteert (hoofdstuk 5). Hier wordt beargumenteerd dat perceptie en cognitie van enerzijds de grafisch-spatiële user interface en anderzijds van de lineair-procedurele gebruiksinstructies verschillend verwerkt worden. Betoogd wordt dat de gebruiker een mentale representatie opbouwt die veeleer lineair-procedureel van aard is (analoog aan de lineair-procedurele gebruiksinstructies) dan grafisch-spatieel (analoog aan de user interface). De mentale representatie kan beter weergegeven worden als een steeds verdere abstractie van de procedures die de gebruiker moet uitvoeren om de verschillende functies van het product in of uit te schakelen, dan door een soort representatie die uitgaat van de grafisch-spatiële user interface. De gebruiker dient veeleer een lineair patroon te herkennen dan een grafisch patroon en het is daarom wellicht van belang om instructies en user interface zoveel mogelijk congruent daarmee te ontwerpen. Er wordt dan ook betoogd om producten zo te ontwerpen dat gebruikers zoveel mogelijk steeds eenzelfde patroon van handelingen moeten uitvoeren ter activering of de-activering van functies.

Hoofdstuk 6 bespreekt de soorten informatie die voor kunnen komen in gebruiksinstructies. Hier wordt de argumentatie van Bieger & Glock gevolgd die stelt dat conclusies betreffende de relatieve efficiëntie van tekst en beeld slechts generaliseerbaar zijn indien ze gebaseerd zijn op specifiek gedefinieerde soorten informatie die in gebruiksinstructies kunnen voorkomen. De negen informatietypen die Bieger & Glock hebben onderscheiden in montage-instructies worden kritisch besproken, maar zullen toch als uitgangspunt dienen. Vervolgens wordt behandeld in hoeverre deze taxonomie toegepast kan worden voor gebruiksinstructies voor producten met een user interface die geheel of nagenoeg geheel tweedimensionaal is, zoals bij moderne elektronische producten en programmatuur.

Dan volgen drie hoofdstukken (7, 8 en 9) waarin de relatieve efficiëntie van tekst, beeld en animatie voor enkele specifieke soorten informatie experimenteel onderzocht wordt. Hoofdstuk 7 behandelt eerst een onderzoek naar de soorten informatie die dominant voorkomen in gebruiksinstructies voor producten met een (nagenoeg) tweedimensionale user interface. Waar Bieger & Glock voor het monteren van driedimensionale producten inventariserende, operationele, spatiële en contextuele informatie van wezenlijk belang vonden, wijst dit onderzoek uit dat, voor producten met een (nagenoeg) tweedimensionale user interface, operationele, spatiële, covariante en contextuele informatie dominant aanwezig zijn. Een test met enkele variaties van de presentatiemedia voor contextuele, operationele en covariante informatie in een complete handleiding voor een bedrijfstelefoon wees uit dat covariante informatie (in dit soort producten meestal de feedback) niet leek bij te dragen tot grotere efficiëntie; ook contextuele informatie (die aangeeft wat het eind- of tussenresultaat na een aantal handelingen moet zijn) leek eerder overbodig dan nuttig. Het vervangen van tekst voor operationele informatie door realistische tekeningen leidde tot relatief efficiënte uitvoering van de taken. Vervangen van tekst door stroomdiagrammen bleek noch voor contextuele informatie, noch voor operationele informatie, noch voor covariante informatie tot efficiënte uitvoering van de taken te leiden.

Hoofdstuk 8 beschrijft een test waarin alleen operationele informatie als instructie aangeboden wordt in tekst, beeld of animatie, al dan niet aangevuld met spatiële informatie in beeld. Ook in deze test dienden proefpersonen een aantal taken uit te voeren op een bedrijfstelefoon met veel functies. Hieruit bleek tekst het meest efficiënte medium en animatie veruit het minst efficiënte medium, zowel wanneer er apart spatiële informatie bij aangeboden werd als wanneer dat niet het geval was. Het aanbieden van deze aparte spatiële informatie in beeld bleek trouwens contra-productief - dit soort informatie kan op deze wijze dus beter niet dan wel aangeboden worden. Het is opvallend dat contextuele informatie (die de gewenste eindtoestand beschrijft) kennelijk niet nodig is voor goed begrip van het systeem. Een vergelijking van de nagenoeg identieke eerste vier taken en laatste vier taken wees uit dat de verschillen in efficiëntie snel afnemen: men leert kennelijk sneller met de beeld- en animatieversies.

In een volgend onderzoek (hoofdstuk 9) is de test beschreven in hoofdstuk 8 herhaald, met als aanvulling dat dezelfde proefpersonen dezelfde taken een week later nogmaals dienden uit te voeren, met dezelfde instructies in hetzelfde medium. Deze test bevestigde de conclusies uit de voorgaande. De herhaling van de opdrachten een week later bevestigde het vermoeden dat animatie voor operationele informatie in eerste instantie een grote tijdsinvestering betekende (inefficiënt), maar dat de meertijd die proefpersonen voor de taken nodig hadden ten opzichte van proefpersonen die deze informatie in tekst of beeld kregen, relatief afnam. Indien de tijd voor het consulteren van de hulp niet meegerekend zou worden in de totale tijd nodig voor het uitvoeren van de taken, is animatie zelfs het meest efficiënte medium bij herhaling van de opdrachten. Herhaling van de nagenoeg identieke laatste vier opdrachten in de tweede test leverde overigens geen verdere verbetering van de efficiëntie op.

Van geheel andere aard is het onderzoek beschreven in hoofdstuk 10. Hierin wordt betoogd dat gebruikers van producten met displays en beeldschermen er voordeel van zouden kunnen hebben als de on-line hulp interactief gemaakt zou worden waar dat mogelijk is. Met 'interactiviteit' wordt bedoeld dat de gebruiker de instructie niet alleen aangeboden krijgt, maar deze ook direct vanuit het hulp-systeem kan uitvoeren. Een on-line hulp-systeem zou hiermee gedeeltelijk de 'gewone' user interface zelfs kunnen vervangen. Dit onderzoek richtte zich op producten met kleine displays, waarbij menu-sturing niet altijd zinvol is; niet gekeken is naar interactiviteit voor programmatuur en apparatuur met grote beeldschermen. Er is van uitgegaan dat het intikken van korte codes efficiënter is dan het werken met menu's indien men de codes uit het hoofd weet. Een vergelijkende test met een aantal verschillende hulp-systemen op een virtuele telefoon (waarbij alle functies op beeldscherm werken) wees uit dat het interactieve hulp-systeem inderdaad leek te leiden tot een efficiëntere uitvoering van taken, bij het bedienen van commando-gestuurde functies, maar het verschil was niet significant. Een menu-gestuurd systeem bleek echter efficiënter, in elk geval voor beginnende gebruikers.

De algemene conclusies (hoofdstuk 11) bestaan voor een deel uit een samenvatting van de conclusies van de afzonderlijke hoofdstukken. Er wordt echter ook een poging gedaan om verdergaande conclusies te trekken uit dit overzicht. Daarbij worden echter ook nogal wat kanttekeningen geplaatst bij de generaliseerbaarheid van de conclusies. Verder wordt geconstateerd dat de taxonomie van Bieger & Glock, die uitgangspunt was voor de tests in de hoofdstukken 7, 8 en 9, enerzijds zeer herkenbare typen informatie in gebruiksinstructies opleverde, maar anderzijds toch ook veel gebreken vertoont. Hoewel de afbeeldingen bij instructies voor tweedimensionale producten veel minder arbitrair kunnen zijn dan die voor driedimensionale producten, is het ook hier niet altijd mogelijk de verschillende soorten informatie geheel onderscheiden visueel aan te bieden.

Vervolgens worden enkele aanbevelingen voor de praktijk gedaan. Vooral op basis van de analyse in hoofdstuk 3 wordt gepleit voor een intensievere samenwerking tussen ontwerpers van de user interface en ontwerpers van de user assistance. Ook de integratie van gebruiksinstructies met pre-sales informatie wordt aanbevolen. Op basis van de argumentatie in hoofdstuk 5 wordt aanbevolen om user

interfaces zo te ontwerpen dat ze zoveel mogelijk aansluiten bij het procedure-model dat gebruikers opbouwen tijdens installatie en bediening van het product. Het verschil tussen het maken van een keuze die leidt tot een optimale manier van handelen en een keuze die leidt tot een voor de gebruiker voldoende efficiënte manier van handelen, wijst naar het advies om te accepteren dat gebruikers vaak tevreden zijn met een minder snelle, maar meer gestandaardiseerde procedures - en daar user interface en user assistance op te ontwerpen. Tevens worden enkele specifieke aanbevelingen gedaan die gericht zijn op het minimaliseren van gebruiksinstructies.
Tenslotte wordt gepoogd te verklaren waarom het niet-congruente medium tekst in eerste instantie veel efficiënter bleek dan het wel congruente medium beeld. Gesuggereerd wordt dat beeld, (voor operationele informatie) cognitief niet verwerkt wordt als beeld en dat dus mentaal een omzetting plaats moet vinden vooraleer men de instructie kan uitvoeren. Dit zou betekenen dat er geen sprake is van 'dual coding' en dat de 'mind's eye hypothese' niet bevestigd wordt. Het is wel een ondersteuning van de Soar-theorie die ervan uitgaat dat mensen alle informatie op één uniforme manier verwerken.

Thanks!

It is not easy to do scientific research that opens up new horizons that also has direct relevance for practice. It is even less easy if the subject deals with such diverse fields of study as psychology, ergonomics, linguistics, communication, industrial design engineering and visual information design. In these fields of study there are sometimes rather different ideas about science and what the study must, should or could imply for daily practice. This is one of the reasons why there was no direct way for the realization of the present thesis: themes appeared and disappeared, the point-of-view for some items changed a couple of times. It is also the reason why the influence of many people, with a wide variety of theoretical and practical backgrounds, is noticeable.

For the realization of this book I should thank everybody I ever met - and many that I have never met. Nevertheless, all responsibility for the final result, especially for all errors, fallacies, follies, megalomania and shortcomings lies purely with me.

Some people have had more influence than others and I will only mention the two who have had most influence on the realization of this thesis. I am very grateful to Paul Mijksenaar for giving me the opportunity to do this study, for the freedom that he has offered me to do this work, for his open mind towards the subject, towards my way of working on it, and towards my character. I also owe many thanks to Hans Dirken for his extensive, critical and stimulating comments, and for his organisational and mental support. In my view our cooperation was very fruitful and instructive. I thank Karel van der Waarde for his extensive reading and for making many useful detailed and general comments. Jans Aasman pinpointed various weak points and exaggerated statements - and I am very glad he did.

Biographical note

Piet Westendorp (Baarn, 1950) studied Dutch language & literature, specializing in rhetoric and argumentation at Utrecht University, worked as a laboratory assistant, copywriter, technical author and journalist for computer and management magazines, and did many other jobs. As a lecturer at Eindhoven University of Technology he has been teaching logic and argumentation, marketing communication and technical communication; at Delft University of Technology he specialized in technical communication and visual information design. He has organized several conferences and published many articles and some books about user instructions and presented various courses and workshops on this subject, both national and international. He has participated in international exhibitions about visual instructions and worked as a consultant and trainer, both freelance and as contract research for national and international companies and (governmental) organizations. He is a general editor of the Information Design Journal.

Stellingen

1 De mentale representatie die een gebruiker van een technisch complex product opbouwt, kan beter gezien worden als een hiërarchisch geordend patroon van procedures op verschillende abstractieniveaus dan als een soort beeldende weergave van user interface of bedieningspaneel.

2 Het kan de voorkeur verdienen om bij het ontwerpen van user interfaces en gebruiksinstructies niet uit te gaan van het 'optimizing model' (waarbij de gebruiker te allen tijde probeert de meest efficiënte manier van handelen te vinden om functies in- of uit te schakelen), maar van het 'satisficing model' (waarbij de gebruiker zich realiseert misschien niet de meest efficiënte manier van handelen toe te passen, maar een manier die zeker tot het gewenste resultaat leidt).

3 Consistentie in bedieningsprocedures is belangrijker voor gebruiksgemak dan variatie en vrijheid in bedieningsmogelijkheden.

4 In de Nederlandse argumentatieleer wordt ten onrechte geen onderscheid gemaakt tussen 'overtuigen' en 'overreden'.

5 Alle aanvullende functies van programmatuur en digitale apparatuur dienen standaard 'uit' te staan. De hulp kan de gebruiker wijzen op het bestaan ervan en op de manier om ze in te schakelen.

6 Het vertalen van [psychologische] theorie naar [vormgevings]praktijk wordt maar al te vaak overgelaten aan afstudeerders [Industrieel Ontwerpen].

7 Naarmate elektronische apparaten intelligenter worden, moeten gebruikers ervan intelligenter zijn om ze te kunnen gebruiken.

8 Grafische gebruikers-interfaces bestaat grotendeels uit taal.

9 De uitspraak 'Het kan niet missen' nadat men iemand de weg heeft gewezen, is vrijwel altijd onjuist.

10 De mobiele telefoon kwam net op tijd om het kwaliteitsverlies van de NS de compenseren.

11 Het woord 'natuurbeheer' is een contradictio in terminis.

Propositions

1 The mental representation that a user of a technical complex product develops is better represented by a hierarchically ordered pattern of procedures in various levels of abstraction than a graphical pattern that has a resemblance to a graphical user interface or display.

2 When designing user interfaces and user instructions, it may be preferable to suppose that users will make satisficing choices, rather than optimizing choices.

3 Consistency in operating procedures is more important for ease of use than variation and freedom of choice in various ways to operate a technical device.

4 It is incorrect that no distinction is made between to 'convince' and to 'persuade' in the Dutch argumentation theory.

5 All extra features in software should be switched 'off' at delivery of the product. The user assistance can help show the existence of such extra features and show how to activate or de-activate them.

6 Transforming [psychological] theory into [design] practice is too often left to the students doing this for their master's thesis [Industrial Design].

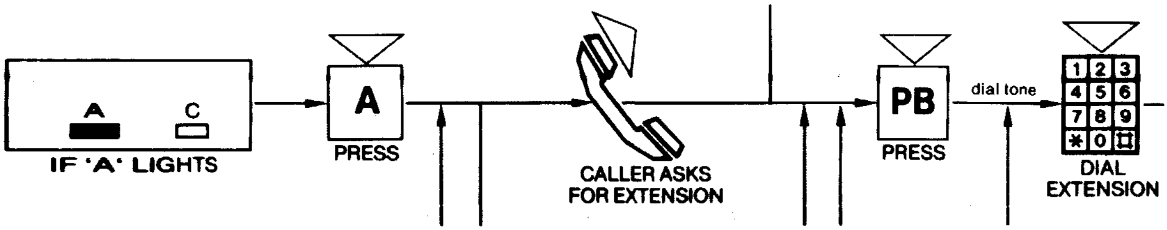
7 The more intelligent electronic devices become, the more intelligent users must be to be able to use these devices.

8 The Graphical User Interface mainly consists of language.

9 The expression 'You can't go wrong' after showing someone the way, is always wrong.

10 The mobile phone arrived just in time to compensate for the decreased quality of the Dutch Railroad company.

11 The literal translation of the Dutch word 'natuurbeheer' is 'nature management', which is a contradiction in terms.


A
C
IF 'A' LIGHTS
A
PRESS
CALLER ASKS
FOR EXTENSION
PB
PRESS
dial tone
1 2 3
4 5 6
7 8 9
* 0 #
DIAL
EXTENSION

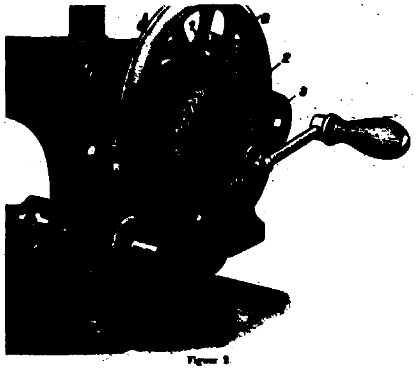


Figure 2

Presentation media for product interaction

Piet Westendorp

Presentation media for product interaction

[addendum]

Proefschrift
ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus, prof. dr. ir. J.T. Fokkema,
voorzitter van het College van Promoties,
in het openbaar te verdedigen op
maandag 10 juni 2002 om 16:00 uur
door

Pieter Hendrikus WESTENDORP
doctorandus in de Nederlandse Taal- en Letterkunde
geboren te Baarn

Dit proefschrift is goedgekeurd door de promotoren:
Prof. P.P. Mijksenaar
Prof. dr. J.M. Dirken Eur Ing.

Samenstelling promotiecommissie:
Rector Magnificus, voorzitter
Prof. P.P. Mijksenaar, Technische Universiteit Delft, promotor
Prof. dr. J.M. Dirken Eur Ing., Technische Universiteit Delft, promotor
Prof. dr. J. Aasman, Technische Universiteit Delft
Prof. dr. D.G. Bouwhuis, Technische Universiteit Eindhoven
Prof. dr. J.W. Drukker, Technische Universiteit Delft
Prof. dr. C. Strothotte, Hochschule Magdeburg, Duitsland
Dr. K.M. van der Waarde, Technische Universiteit Delft, adviseur

ISBN 90-9015900-2

Cover illustrations

Cataract operation. 12th Century. Manuscript Sloane. London Library (Herrlinger, 1967, p. 17).

Do not use lighters. US Air - Boeing 757 Safety card

Arrow on sewing machine. Naumann KL-14, late 19th century.

Philips TBX telephone operator manual. Flowchart.

Presentation media for product interaction

Summary

This study focuses on the relative efficiency of text, pictures and animation for some specific types of information in user instructions. Chapter 1 presents a general introduction with a motivation from theory and from practice for this study. Chapter 2 presents an overview of the increasing necessity, historic development and design aspects of user instructions. Here it is argued that technical (consumer) products are provided with more and more features, and are therefore rapidly becoming more complex and difficult to use. As a result, user instructions are becoming more important, in spite of all efforts to improve user interfaces.

The overview in chapter 2 shows that specific user instructions (for a specific type and model of a product) originated around the middle of the 19th century, when mass production of complex technical consumer products started. Both conceptual designs and detailing of the user instructions develop on the basis of generic instruction books that have been produced since the Middle Ages, especially for medical, technical and military procedures. New developments in visual instructions are discussed, including instructive elements (such as arrows, indicating hands and color for indicating parts). Conclusions for practical application about the relative efficiency of text, pictures and animation for user instructions must be judged within this context of possible concepts and detailing.

In the next three chapters (3, 4 and 5) the communication and interaction process between user and technical product is discussed; first on a sociological level (relation between producers and users of instructions), then on a meso-level (communication between the user of a product, the product and the user instructions), and finally on a psychological level (perception and cognition). The starting point here is that the user communicates and interacts with the products' user interface and with the supporting user assistance as well.

Chapter 3 describes the complex communication process between designers and users of user instructions. This chapter shows the overwhelming number of factors involved in each part of the process: making, transmitting and consulting the message. User instructions are not made by individuals, but by teams of professionals with a wide variety of backgrounds, such as lawyers, technicians, programmers, graphic designers, writers and illustrators. The target group may consist of hundreds of millions of people with a very wide variety of backgrounds, educations, technical inclination, etc. User assistance may consist of many different elements, such as advertisements, user instructions, on-line help systems and guided tours. Often the user will have made some kind of mental representation of a product before he bought it. Moreover, the 'message' in this process is often very complex: especially electronic products may have very many features. Possible practical conclusions of studies contrasting the relative efficiency of various media for instruction must also take these kinds of factors into account.

Chapter 4 zooms in on the communication and interaction process between user, product and user instructions. Too often the focus has been on just the relation between user and product (user interface). First the focus is on how often the users have to consult both the user interface and the user instructions for every detailed step in each procedure of the user instructions. Next, a model is developed which describes this switching between user interface and user instructions. The goal of this model is to map all user tasks in this triangle between the user and user instructions and the product.

This switching between user interface and user instructions is the starting point for a rather extensive discussion of the perception and cognitive processing of the interaction and communication of someone who tries to operate a complex technical product while consulting the user instructions (chapter 5). It is argued that perception and cognition processing of the graphical-spatial user interface, on the one hand, and the procedural-linear user instructions, on the other hand, could be quite different. It is further argued that users may develop mental representations that are rather more procedural-linear (analogue to the user instructions) than graphical-spatial (analogue to the user interface). If so, the mental representation is better depicted as an ongoing abstraction of procedures that users have to follow to activate or de-activate features of the product than a depiction based on the graphical-spatial user interface. Users have to recognize a linear pattern, rather than a graphical pattern, which is why it might be useful to try and design user instructions and user interfaces as congruent as possible. As a result of this reasoning, it is argued that user interface should be designed so that users can follow the same procedures for activating and de-activating features as much as possible.

Chapter 6 discusses the types of information that may occur in user instructions. From the argumentation by Bieger & Glock it follows that conclusions from tests contrasting the relative efficiency of various media for instructions can only be generalized if they are based on specified types of information in user instructions. The nine types of information that Bieger & Glock have discerned in assembling instructions are criticized, but yet they are used as a basis for tests. Then a discussion follows about the possibility to apply the Bieger & Glock taxonomy to user instructions for products with two-dimensional or nearly two-dimensional user interfaces, such as modern electronic products or software.

The next three chapters (7, 8 and 9) report about tests contrasting various presentation media for some specified types of information. In chapter 7, first the types of information that are dominant in user instructions for products with a (nearly) two-dimensional user interface are discussed. Where Bieger & Glock found inventory, operational, spatial and contextual information to be dominant (for assembly instructions), in this study operational, spatial, covariant and contextual information appeared to be dominant. A test with some variations of the presentation media for operational, contextual and covariant information in a complete manual for a telephone system indicated that covariant information (in these products usually feedback information) did not contribute to the efficiency. Contextual information (usually indicating what should be the result of a series of actions in a procedure) also seemed counter-productive rather than useful. Replacing text for the presentation of operational information with representational pictures led to a relatively efficient completion of tasks by subjects in the test. Replacing texts with flowcharts appeared not to be efficient for operational, nor for contextual nor for covariant information.

Chapter 8 presents a test in which either only operational information is presented to subjects who had to perform tasks on a telephone, or operational and spatial information. The operational information was presented in either text, pictures or animations; the spatial information only in separate pictures. The results indicate that text is the most efficient medium and animation by far the least efficient medium, both with or without the added spatial information in separate pictures. Presenting spatial information in separate pictures appeared to be very inefficient - subjects performed better with the versions without this information. The results also indicated that contextual information appears not to be required for a good understanding of the system; subjects performed very well without this information. A comparison of the first four tasks that subjects had to do and the nearly identical last four tasks indicated that the differences between the relative efficiency of text, pictures and animations had become much smaller within the same test that lasted at most half an hour. It seems that subjects had learned faster with the picture and animation versions.

The next chapter (9) describes a test that is identical to the one discussed in chapter 8, but in this test the same subjects were asked to repeat the same tasks using the same instructions in the same medium one week later. This test confirmed the conclusions of the previous test. Repetition of the tasks one week later showed that animation indeed proved to be more efficient on repetitive use of the product. Animation for operational information requires a big investment of time (compared to text and pictures), but it pays off with repetitive application of the same tasks. The time subjects needed for completion of the tasks decreased more than for subjects using instructions in other media. If the time needed just for consultation of the instructions were not taken into account, animations would be the most efficient medium when the tasks would have to be repeated. However, repetition of the nearly identical tasks at the end of the second test did not show any further improvement.

A totally different study is presented in chapter 10. Here it is argued that interactive on-line help could be beneficial for users of products with displays and screens. 'Interactivity' here means that the on-line help system does not just offer instructions, but also allows the user to immediately perform the instructions from within the help system. Thus, the on-line help would partially replace the 'regular' user interface. This study focused on products with small displays, where guidance via menus may not always be useful; products and software with screens were not included. It was presupposed that entering short codes would be more efficient if the user knew these codes by heart. A contrastive test with variations of a help system for a virtual telephone (with all functions working realistically) indicated that an interactive help system could indeed lead to a more efficient performance of the tasks for products when codes had to be entered in a command-line interface, but the difference was not significant. However, a menu-driven system appeared more efficient, at least for first time users.

The general conclusions are presented in chapter 11, with some remarks concerning the generalizability. It is further concluded that the Bieger & Glock taxonomy, which was the basis for the tests in chapters 7, 8 and 9, on the one hand produced very well recognizable types of information in user instructions, but on the other hand was also subject to severe criticism. Although the pictures used for instructions for products with a two-dimensional user interface can be far less arbitrarily chosen than pictures for assembly instructions, it was not always possible to fully separate the several types of information and offer just one type pictorially.

Next, some suggestions for practice are discussed. Especially on the basis of chapter 3, a plea is made for a more intense collaboration between user interface designers and user assistance designers. A further integration between user instructions and pre-sales information (such as advertising) is also made. On the basis of the argumentation in chapter 5, it is advised to design user interfaces to correspond as much as possible with the procedure model that users may develop during installation and use of the product. The difference between the optimizing choice and the 'satisficing' choice leads to the advice to accept that users are often content with a less efficient way to operate the product, if it is more standard, and to possibly design user interface and user instructions on this basis. Then some specific advice is presented concerning minimalization of manuals.

Finally an attempt is made to explain why the incongruent medium 'text' is more efficient during first use than the more congruent medium 'pictures'. It is suggested that pictures (for operational information in user instructions) are cognitively not processed as pictures, but have to be converted to a more abstract representation before users can execute the instruction. This would imply that information is not processed in a dual coded way, and that the 'mind's eye hypothesis cannot be confirmed. If this is the case, it would support the Soar theory, which assumes that humans process all information in one unitary way.

Presentatiemedia voor interactie met producten

Samenvatting [in Dutch]

Dit boek behandelt de relatieve efficiëntie van tekst, beeld en animatie voor enkele specifieke soorten informatie die voorkomen in gebruiksinstructies. Daartoe wordt - na de algemene introductie in hoofdstuk 1- eerst in hoofdstuk 2 een overzicht gegeven van de noodzaak, de historische ontwikkeling en de ontwerpaspecten van gebruiksinstructies. Beargumenteerd wordt dat de technische (consumenten)producten snel complexer worden omdat ze steeds meer mogelijkheden krijgen. Daarom worden gebruiksinstructies steeds belangrijker, ondanks pogingen om user interfaces van producten gebruiksvriendelijker te maken. Het overzicht laat zien dat specifieke gebruiksinstructies voor een speciaal merk en type product rond het midden van de negentiende eeuw ontstonden met de komst van seriematig geproduceerde technische apparaten voor een groot publiek. Zowel concepten als detailleringen hiervan vloeien voort uit de generieke instructies die sinds de middeleeuwen zijn gemaakt, vooral voor medische, technische en militaire handelingen. De nieuwe ontwikkelingen op het gebied van de visuele instructies worden behandeld, waaronder de instructieve elementen (zoals pijlen, handen die aanwijzen en gebruik van kleur om een deel aan te duiden). Eventuele op de praktijk gerichte conclusies ten aanzien van de relatieve efficiëntie van de verschillende media (tekst, beeld en animatie) voor gebruiksinstructies dienen gezien te worden binnen deze context van mogelijke concepten en detaillering.

Vervolgens wordt in een drietal hoofdstukken (3, 4, 5) het communicatie- en interactieproces tussen gebruiker en technisch product behandeld; eerst op sociologisch niveau (communicatie tussen makers en gebruikers van instructies), dan op meso-niveau (interactie van gebruiker met product en gebruiksinstructie) en vervolgens op psychologisch niveau ('het denken van den gebruiker'). Hierbij dient als uitgangspunt dat de gebruiker enerzijds communiceert en interacteert met de user interface van het product en anderzijds met de

user assistance, de aanvullende ondersteuning. Eerst (hoofdstuk 3) wordt het communicatieproces tussen producent en gebruiker van instructies beschreven. Dit hoofdstuk toont de hoeveelheid factoren die een rol spelen bij het maken, communiceren en consulteren van de gebruiksinstructies. De 'maker' is een team dat kan bestaan uit juristen, technici, programmeurs, schrijvers, tekenaars en vele anderen. De doelgroep kan uit honderden miljoenen mensen te bestaan, van velerlei achtergrond, opleiding, belangstelling, technische interesse, enzovoort. De assistentie voor de gebruiker kan bestaan uit een groot aantal middelen, zoals advertenties, de gebruiksaanwijzing, het on-line help-systeem en de 'rondleiding' (waarbij de meest interessante kenmerken van het product getoond worden). Bovendien is de 'boodschap' in dit proces vaak uitermate complex: met name elektronische producten en programmatuur kunnen honderden aanvullende functies te hebben. Eventuele praktische conclusies van onderzoek naar de relatieve efficiëntie van tekst, beeld en animatie voor instructie dienen ook rekening te houden met al deze factoren.

In hoofdstuk 4 wordt verder 'ingezoomd' op het communicatie- en interactieproces tussen gebruiker, product en gebruiksinstructies. Vaak wordt bij dit proces alleen acht geslagen op de gebruiker en het product. Hier wordt eerst beschreven hoe gebruikers van technisch complexe producten bij elke detailstap in een instructieprocedure zowel user interface als gebruiksinstructie moeten consulteren. Vervolgens wordt een model ontwikkeld dat dit 'heen-en-weer schakelen' tussen de user interface van het product en de gebruiksinstructie probeert te beschrijven. Het model heeft als doel alle taken van de gebruiker in deze driehoeksverhouding in hun onderling verband onder te brengen.

Dit 'heen-en-weer schakelen' tussen user interface en gebruiksinstructies vormt het uitgangspunt voor een beschouwing van perceptie en cognitieve verwerking van de interactie en communicatie van iemand die een technisch complex product probeert te gebruiken en daarbij de gebruiksinstructies consulteert (hoofdstuk 5). Hier wordt beargumenteerd dat perceptie en cognitie van enerzijds de grafisch-spatiële user interface en anderzijds van de lineair-procedurele gebruiksinstructies verschillend verwerkt worden. Betoogd wordt dat de gebruiker een mentale representatie opbouwt die veeleer lineair-procedureel van aard is (analoog aan de lineair-procedurele gebruiksinstructies) dan grafisch-spatieel (analoog aan de user interface). De mentale representatie kan beter weergegeven worden als een steeds verdere abstractie van de procedures die de gebruiker moet uitvoeren om de verschillende functies van het product in of uit te schakelen, dan door een soort representatie die uitgaat van de grafisch-spatiële user interface. De gebruiker dient veeleer een lineair patroon te herkennen dan een grafisch patroon en het is daarom wellicht van belang om instructies en user interface zoveel mogelijk congruent daarmee te ontwerpen. Er wordt dan ook betoogd om producten zo te ontwerpen dat gebruikers zoveel mogelijk steeds eenzelfde patroon van handelingen moeten uitvoeren ter activering of de-activering van functies.

Hoofdstuk 6 bespreekt de soorten informatie die voor kunnen komen in gebruiksinstructies. Hier wordt de argumentatie van Bieger & Glock gevolgd die stelt dat conclusies betreffende de relatieve efficiëntie van tekst en beeld slechts generaliseerbaar zijn indien ze gebaseerd zijn op specifiek gedefinieerde soorten informatie die in gebruiksinstructies kunnen voorkomen. De negen informatietypen die Bieger & Glock hebben onderscheiden in montage-instructies worden kritisch besproken, maar zullen toch als uitgangspunt dienen. Vervolgens wordt behandeld in hoeverre deze taxonomie toegepast kan worden voor gebruiksinstructies voor producten met een user interface die geheel of nagenoeg geheel tweedimensionaal is, zoals bij moderne elektronische producten en programmatuur.

Dan volgen drie hoofdstukken (7, 8 en 9) waarin de relatieve efficiëntie van tekst, beeld en animatie voor enkele specifieke soorten informatie experimenteel onderzocht wordt. Hoofdstuk 7 behandelt eerst een onderzoek naar de soorten informatie die dominant voorkomen in gebruiksinstructies voor producten met een (nagenoeg) tweedimensionale user interface. Waar Bieger & Glock voor het monteren van driedimensionale producten inventariserende, operationele, spatiële en contextuele informatie van wezenlijk belang vonden, wijst dit onderzoek uit dat, voor producten met een (nagenoeg) tweedimensionale user interface, operationele, spatiële, covariante en contextuele informatie dominant aanwezig zijn. Een test met enkele variaties van de presentatiemedia voor contextuele, operationele en covariante informatie in een complete handleiding voor een bedrijfstelefoon wees uit dat covariante informatie (in dit soort producten meestal de feedback) niet leek bij te dragen tot grotere efficiëntie; ook contextuele informatie (die aangeeft wat het eind- of tussenresultaat na een aantal handelingen moet zijn) leek eerder overbodig dan nuttig. Het vervangen van tekst voor operationele informatie door realistische tekeningen leidde tot relatief efficiënte uitvoering van de taken. Vervangen van tekst door stroomdiagrammen bleek noch voor contextuele informatie, noch voor operationele informatie, noch voor covariante informatie tot efficiënte uitvoering van de taken te leiden.

Hoofdstuk 8 beschrijft een test waarin alleen operationele informatie als instructie aangeboden wordt in tekst, beeld of animatie, al dan niet aangevuld met spatiële informatie in beeld. Ook in deze test dienden proefpersonen een aantal taken uit te voeren op een bedrijfstelefoon met veel functies. Hieruit bleek tekst het meest efficiënte medium en animatie veruit het minst efficiënte medium, zowel wanneer er apart spatiële informatie bij aangeboden werd als wanneer dat niet het geval was. Het aanbieden van deze aparte spatiële informatie in beeld bleek trouwens contra-productief - dit soort informatie kan op deze wijze dus beter niet dan wel aangeboden worden. Het is opvallend dat contextuele informatie (die de gewenste eindtoestand beschrijft) kennelijk niet nodig is voor goed begrip van het systeem. Een vergelijking van de nagenoeg identieke eerste vier taken en laatste vier taken wees uit dat de verschillen in efficiëntie snel afnemen: men leert kennelijk sneller met de beeld- en animatieversies.

In een volgend onderzoek (hoofdstuk 9) is de test beschreven in hoofdstuk 8 herhaald, met als aanvulling dat dezelfde proefpersonen dezelfde taken een week later nogmaals dienden uit te voeren, met dezelfde instructies in hetzelfde medium. Deze test bevestigde de conclusies uit de voorgaande. De herhaling van de opdrachten een week later bevestigde het vermoeden dat animatie voor operationele informatie in eerste instantie een grote tijdsinvestering betekende (inefficiënt), maar dat de meertijd die proefpersonen voor de taken nodig hadden ten opzichte van proefpersonen die deze informatie in tekst of beeld kregen, relatief afnam. Indien de tijd voor het consulteren van de hulp niet meegerekend zou worden in de totale tijd nodig voor het uitvoeren van de taken, is animatie zelfs het meest efficiënte medium bij herhaling van de opdrachten. Herhaling van de nagenoeg identieke laatste vier opdrachten in de tweede test leverde overigens geen verdere verbetering van de efficiëntie op.

Van geheel andere aard is het onderzoek beschreven in hoofdstuk 10. Hierin wordt betoogd dat gebruikers van producten met displays en beeldschermen er voordeel van zouden kunnen hebben als de on-line hulp interactief gemaakt zou worden waar dat mogelijk is. Met 'interactiviteit' wordt bedoeld dat de gebruiker de instructie niet alleen aangeboden krijgt, maar deze ook direct vanuit het hulp-systeem kan uitvoeren. Een on-line hulp-systeem zou hiermee gedeeltelijk de 'gewone' user interface zelfs kunnen vervangen. Dit onderzoek richtte zich op producten met kleine displays, waarbij menu-sturing niet altijd zinvol is; niet gekeken is naar interactiviteit voor programmatuur en apparatuur met grote beeldschermen. Er is van uitgegaan dat het intikken van korte codes efficiënter is dan het werken met menu's indien men de codes uit het hoofd weet. Een vergelijkende test met een aantal verschillende hulp-systemen op een virtuele telefoon (waarbij alle functies op beeldscherm werken) wees uit dat het interactieve hulp-systeem inderdaad leek te leiden tot een efficiëntere uitvoering van taken, bij het bedienen van commando-gestuurde functies, maar het verschil was niet significant. Een menu-gestuurd systeem bleek echter efficiënter, in elk geval voor beginnende gebruikers.

De algemene conclusies (hoofdstuk 11) bestaan voor een deel uit een samenvatting van de conclusies van de afzonderlijke hoofdstukken. Er wordt echter ook een poging gedaan om verdergaande conclusies te trekken uit dit overzicht. Daarbij worden echter ook nogal wat kanttekeningen geplaatst bij de generaliseerbaarheid van de conclusies. Verder wordt geconstateerd dat de taxonomie van Bieger & Glock, die uitgangspunt was voor de tests in de hoofdstukken 7, 8 en 9, enerzijds zeer herkenbare typen informatie in gebruiksinstructies opleverde, maar anderzijds toch ook veel gebreken vertoont. Hoewel de afbeeldingen bij instructies voor tweedimensionale producten veel minder arbitrair kunnen zijn dan die voor driedimensionale producten, is het ook hier niet altijd mogelijk de verschillende soorten informatie geheel onderscheiden visueel aan te bieden.

Vervolgens worden enkele aanbevelingen voor de praktijk gedaan. Vooral op basis van de analyse in hoofdstuk 3 wordt gepleit voor een intensievere samenwerking tussen ontwerpers van de user interface en ontwerpers van de user assistance. Ook de integratie van gebruiksinstructies met pre-sales informatie wordt aanbevolen. Op basis van de argumentatie in hoofdstuk 5 wordt aanbevolen om user

interfaces zo te ontwerpen dat ze zoveel mogelijk aansluiten bij het procedure-model dat gebruikers opbouwen tijdens installatie en bediening van het product. Het verschil tussen het maken van een keuze die leidt tot een optimale manier van handelen en een keuze die leidt tot een voor de gebruiker voldoende efficiënte manier van handelen, wijst naar het advies om te accepteren dat gebruikers vaak tevreden zijn met een minder snelle, maar meer gestandaardiseerde procedures - en daar user interface en user assistance op te ontwerpen. Tevens worden enkele specifieke aanbevelingen gedaan die gericht zijn op het minimaliseren van gebruiksinstructies.
Tenslotte wordt gepoogd te verklaren waarom het niet-congruente medium tekst in eerste instantie veel efficiënter bleek dan het wel congruente medium beeld. Gesuggereerd wordt dat beeld, (voor operationele informatie) cognitief niet verwerkt wordt als beeld en dat dus mentaal een omzetting plaats moet vinden vooraleer men de instructie kan uitvoeren. Dit zou betekenen dat er geen sprake is van 'dual coding' en dat de 'mind's eye hypothese' niet bevestigd wordt. Het is wel een ondersteuning van de Soar-theorie die ervan uitgaat dat mensen alle informatie op één uniforme manier verwerken.

Thanks!

It is not easy to do scientific research that opens up new horizons that also has direct relevance for practice. It is even less easy if the subject deals with such diverse fields of study as psychology, ergonomics, linguistics, communication, industrial design engineering and visual information design. In these fields of study there are sometimes rather different ideas about science and what the study must, should or could imply for daily practice. This is one of the reasons why there was no direct way for the realization of the present thesis: themes appeared and disappeared, the point-of-view for some items changed a couple of times. It is also the reason why the influence of many people, with a wide variety of theoretical and practical backgrounds, is noticeable.

For the realization of this book I should thank everybody I ever met - and many that I have never met. Nevertheless, all responsibility for the final result, especially for all errors, fallacies, follies, megalomania and shortcomings lies purely with me.

Some people have had more influence than others and I will only mention the two who have had most influence on the realization of this thesis. I am very grateful to Paul Mijksenaar for giving me the opportunity to do this study, for the freedom that he has offered me to do this work, for his open mind towards the subject, towards my way of working on it, and towards my character. I also owe many thanks to Hans Dirken for his extensive, critical and stimulating comments, and for his organisational and mental support. In my view our cooperation was very fruitful and instructive. I thank Karel van der Waarde for his extensive reading and for making many useful detailed and general comments. Jans Aasman pinpointed various weak points and exaggerated statements - and I am very glad he did.

Biographical note

Piet Westendorp (Baarn, 1950) studied Dutch language & literature, specializing in rhetoric and argumentation at Utrecht University, worked as a laboratory assistant, copywriter, technical author and journalist for computer and management magazines, and did many other jobs. As a lecturer at Eindhoven University of Technology he has been teaching logic and argumentation, marketing communication and technical communication; at Delft University of Technology he specialized in technical communication and visual information design. He has organized several conferences and published many articles and some books about user instructions and presented various courses and workshops on this subject, both national and international. He has participated in international exhibitions about visual instructions and worked as a consultant and trainer, both freelance and as contract research for national and international companies and (governmental) organizations. He is a general editor of the Information Design Journal.

Stellingen

1 De mentale representatie die een gebruiker van een technisch complex product opbouwt, kan beter gezien worden als een hiërarchisch geordend patroon van procedures op verschillende abstractieniveaus dan als een soort beeldende weergave van user interface of bedieningspaneel.

2 Het kan de voorkeur verdienen om bij het ontwerpen van user interfaces en gebruiksinstructies niet uit te gaan van het 'optimizing model' (waarbij de gebruiker te allen tijde probeert de meest efficiënte manier van handelen te vinden om functies in- of uit te schakelen), maar van het 'satisficing model' (waarbij de gebruiker zich realiseert misschien niet de meest efficiënte manier van handelen toe te passen, maar een manier die zeker tot het gewenste resultaat leidt).

3 Consistentie in bedieningsprocedures is belangrijker voor gebruiksgemak dan variatie en vrijheid in bedieningsmogelijkheden.

4 In de Nederlandse argumentatieleer wordt ten onrechte geen onderscheid gemaakt tussen 'overtuigen' en 'overreden'.

5 Alle aanvullende functies van programmatuur en digitale apparatuur dienen standaard 'uit' te staan. De hulp kan de gebruiker wijzen op het bestaan ervan en op de manier om ze in te schakelen.

6 Het vertalen van [psychologische] theorie naar [vormgevings]praktijk wordt maar al te vaak overgelaten aan afstudeerders [Industrieel Ontwerpen].

7 Naarmate elektronische apparaten intelligenter worden, moeten gebruikers ervan intelligenter zijn om ze te kunnen gebruiken.

8 Grafische gebruikers-interfaces bestaat grotendeels uit taal.

9 De uitspraak 'Het kan niet missen' nadat men iemand de weg heeft gewezen, is vrijwel altijd onjuist.

10 De mobiele telefoon kwam net op tijd om het kwaliteitsverlies van de NS de compenseren.

11 Het woord 'natuurbeheer' is een contradictio in terminis.

Propositions

1 The mental representation that a user of a technical complex product develops is better represented by a hierarchically ordered pattern of procedures in various levels of abstraction than a graphical pattern that has a resemblance to a graphical user interface or display.

2 When designing user interfaces and user instructions, it may be preferable to suppose that users will make satisficing choices, rather than optimizing choices.

3 Consistency in operating procedures is more important for ease of use than variation and freedom of choice in various ways to operate a technical device.

4 It is incorrect that no distinction is made between to 'convince' and to 'persuade' in the Dutch argumentation theory.

5 All extra features in software should be switched 'off' at delivery of the product. The user assistance can help show the existence of such extra features and show how to activate or de-activate them.

6 Transforming [psychological] theory into [design] practice is too often left to the students doing this for their master's thesis [Industrial Design].

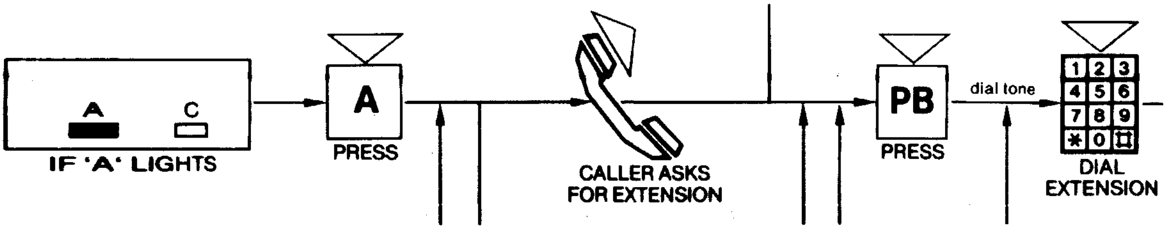
7 The more intelligent electronic devices become, the more intelligent users must be to be able to use these devices.

8 The Graphical User Interface mainly consists of language.

9 The expression 'You can't go wrong' after showing someone the way, is always wrong.

10 The mobile phone arrived just in time to compensate for the decreased quality of the Dutch Railroad company.

11 The literal translation of the Dutch word 'natuurbeheer' is 'nature management', which is a contradiction in terms.


A
C
IF 'A' LIGHTS
A
PRESS
CALLER ASKS
FOR EXTENSION
PB
PRESS
dial tone
1 2 3
4 5 6
7 8 9
* 0
DIAL
EXTENSION

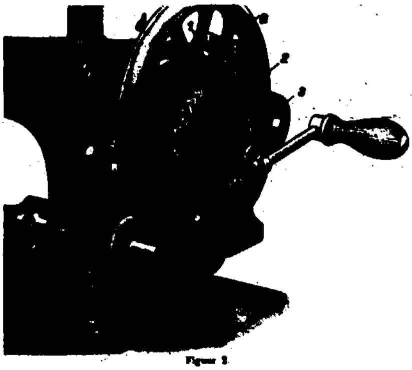


Figure 2